Naturalism

A critical analysis

Edited by
William Lane Craig
and J.P. Moreland

Routledge Studies in Twentieth-Century Philosophy

Also available as a printed book
see title verso for ISBN details

Naturalism

Naturalism: A Critical Analysis presents a rigorous analysis and critique of the major varieties of contemporary philosophical naturalism. It advocates the thesis that contemporay naturalism should be abandoned in light of the serious difficulties raised against it. The contributors draw on a wide range of topics including: epistemology, philosophy of science, value theory to basic analytic ontology, philosophy of mind and agency, and natural theology.

For some time now, philosophical naturalism has played a dominant intellectual and sociological role in academe. Currently, throughout a large sector of professional philosophy, philosophical naturalism determines the way philosophy is to be done, what counts as good methodology, and what is taken to be a worthwhile topic for study. This book provides a thorough and searching critique of philosophical naturalism which is squarely within the analytic tradition and which avails itself of the new and exciting literature in philosophy of religion that has burst onto the scene in the last decade.

Naturalism: A Critical Analysis accomplishes three things. First, it argues that naturalism fails to deal adequately with a number of desiderata; second, it shows that to the degree that naturalism is a consistent paradigm and is 'located' within the physicalist story, it should be a form of strict physicalism; and third, it provides an account of the contemporary resurgence of philosophical theism by advertising to evidentiary considerations in the natural world itself (e.g. Big Bang cosmology) which serve as signposts of transcendence.

William Lane Craig is Research Professor of Philosophy and **J. P. Moreland** is Professor of Philosophy both at Biola University, La Mirada, California.

Routledge Studies in Twentieth-century Philosophy

1 The Story of Analytic Philosophy
Plot and heroes
Edited by Anat Biletzki and Anat Matar

2 Donald Davidson
Truth, meaning and knowledge
Edited by Urszula M. Zeglén

3 Philosophy and Ordinary Language
The bent and genuis of our tongue
Oswald Hanfling

4 The Subject in Question
Sartre's critique of Husserl in *The Transcendence of the Ego*
Stephen Priest

5 Naturalism
A critical analysis
Edited by William Lane Craig and J. P. Moreland

Naturalism

A critical analysis

Edited by
William Lane Craig and
J. P. Moreland

London and New York

First published 2000
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

Routledge is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2001.

Editorial selection and material © 2000 Wiliam Lane Craig and J. P. Moreland

Individual chapters © individual contributors

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication Data
Naturalism: a critical analysis/edited by William Lane Craig and J. P. Moreland.
p. cm.
Includes bibliographical references and index.
1. Naturalism. I. Craig, William Lane. II. Moreland, James Porter, 1948–

B828.2.N367 2000
146—dc21 00–025324

ISBN 0–415–23524–3 (Print Edition)
ISBN 0-203-18613-3 Master e-book ISBN
ISBN 0-203-18736-9 (Glassbook Format)

Contents

List of figures vii
List of contributors ix
Preface xi

PART 1
Epistemology **1**

1 Farewell to philosophical naturalism 3
PAUL K. MOSER AND DAVID YANDELL

2 Knowledge and naturalism 24
DALLAS WILLARD

3 The incompatibility of naturalism and scientific realism 49
ROBERT C. KOONS

PART 2
Ontology **65**

4 Naturalism and the ontological status of properties 67
J.P. MORELAND

5 Naturalism and material objects 110
MICHAEL REA

6 Naturalism and the mind 133
CHARLES TALIAFERRO

7 Naturalism and libertarian agency 156
STEWART GOETZ

PART 3
Value theory **187**

8 Naturalism and morality 189
JOHN E. HARE

PART 4
Natural theology **213**

9 Naturalism and cosmology 215
WILLIAM LANE CRAIG

10 Naturalism and design 253
WILLIAM DEMBSKI

Index 281

Figures

9.1	Conical representation of Standard model space-time	218
9.2	Steady State model	219
9.3	Oscillating model	221
9.4	Oscillating model with entropy increase	222
9.5	Vacuum Fluctuation models	224
9.6	Chaotic Inflationary model	225
9.7	Quantum Gravity model	227
9.8	Contrast between the universe as we know it (assumed for convenience to be closed) with a more probable universe	230
9.9	A self-creating universe	233

Contributors

William Lane Craig is Research Professor of Philosophy at Talbot School of Theology, Biola University, La Mirada, California. He is the author of a dozen books, including *The Kalam Cosmological Argument* and *Theism, Atheism, and Big Bang Cosmology* (with Quentin Smith), as well as over a hundred articles in professional journals such as *The Journal of Philosophy, Philosophy, Philosophical Studies, Australasian Journal of Philosophy, The British Journal for the Philosophy of Science, Erkenntnis, International Studies in the Philosophy of Science* and *Philosophia naturalis.*

William Dembski is a Fellow of Discovery Institute's Center for the Renewal of Science and Culture, an Associate Research Professor at Baylor University, and director of Baylor's Michael Polanyi Center, a research group focusing on complexity, information, design, and the conceptual foundations of science. He has authored and edited several books including *The Design Inference* (Cambridge University Press, 1998). He has also authored over forty articles in such varied journals *as Journal of Theoretical Probability, Nous* and *Scottish Journal of Theology.*

Stewart Goetz is Associate Professor of Philosophy at Ursinus College in Collegeville, PA. He specializes in the philosophy of action and the philosophy of mind, and has authored a number of articles on issues in these areas in journals such as *American Philosophical Quarterly, Faith and Philosophy, Mind, Philosophical Studies* and *Philosophy and Phenomenological Research.*

John E. Hare is Professor of Philosophy at Calvin College in Grand Rapids, Michigan. He has taught at Lehigh University and was a visiting professor at the University of Michigan. His articles have appeared in a number of journals including *Ancient Philosophy, International Journal of Applied Ethics,* and *Phronesis.* He is also the author of *The Moral Gap* (Clarendon, 1996).

Robert C. Koons is an Associate Professor of Philosophy at the University of Texas at Austin. He is the author of *Paradoxes of Belief and Strategic Rationality* (Cambridge University Press, 1992), "A New Look at the Cosmological Argument", *American*

Philosophical Quarterly (1997), and *Realism Regained: An Exact Theory of Causation, Teleology and the Mind* (Oxford University Press, 2000).

J. P. Moreland is professor of philosophy at Talbot School of Theology, Biola University, La Mirada, California. He has authored, co-authored, or edited twelve books including *Does God Exist?* (Prometheus, 1993). He has also written over forty articles in journals that include *Philosophy and Phenomenological Research, American Philosophical Quarterly, Southern Journal of Philosophy, MetaPhilosophy, Religious Studies* and *Australasian Journal of Philosophy*.

Paul K. Moser is Professor and Chairperson of Philosophy at Loyola University of Chicago. He is the author of *Philosophy after Objectivity* (Oxford University Press, 1993) and *Knowledge and Evidence* (Cambridge University Press, 1989), and co-author of *The Theory of Knowledge* (Oxford University Press, 1998). He is also General Editor of the book series, Routledge Contemporary Introductions to Philosophy.

Michael Rea received his Ph.D. in 1996 from the University of Notre Dame. He is currently Assistant Professor of Philosophy at the University of Delaware. His articles on the metaphysics of material objects have appeared in *The Philosophical Review, Philosophy and Phenomenological Research, Philosophical Studies* and *American Philosophical Quarterly*. He is the editor of *Material Constitution: A Reader* (Lanham, MD: Rowman & Littlefield, 1997).

Charles Taliaferro is Professor of Philosophy at St Olaf College, Northfield, Minnesota. He is the author of numerous articles in professional journals. He is also the author of *Consciousness and the Mind of God* (Cambridge, 1994), *Contemporary Philosophy of Religion* (Blackwell, 1998), and co-editor of *A Companion to Philosophy of Religion* (co-edited with Philip L. Quinn) (Blackwell, 1997).

Dallas Willard is Professor of Philosophy at the University of Southern California, where he has been since 1965. He began his teaching career at the University of Wisconsin (Madison), and has held visiting appointments at the University of California at Los Angeles and the University of Colorado. He is the author of *Logic and the Objectivity of Knowledge* and of numerous articles, including, recently, "Knowledge," in *The Cambridge Companion to Husserl*, "Who Needs Brentano? The Wasteland of Philosophy Without Its Past," *The Brentano Puzzle*, and "Degradation of Logical Form," in *Axiomathes* (1997).

David Yandell is Associate Professor of Philosophy at Loyola University of Chicago. He is the author of a number of recent articles in the history of philosophy which have appeared in such journals as *British Journal for the History of Philosophy* and *History of Philosophy Quarterly*.

Preface

Throughout much of contemporary intellectual culture, the world-view of naturalism is the accepted creed. David Papineau boldly proclaims that ". . . nearly everybody nowadays wants to be a 'naturalist' . . ."[1] Along similar lines, Jaegwon Kim states that most philosophers today have accepted naturalism – and more specifically, physicalism, as the best contemporary version of naturalism – as a basic commitment. Such a commitment, says Kim, involves recognizing that naturalism "is imperialistic; it demands 'full coverage' . . . and exacts a terribly high ontological price."[2] This volume aims to challenge this prevailing world-view by making clear the price which must be paid to accept it.

Both in its various historical and contemporary manifestations, it would be hard to state naturalism to everyone's satisfaction. Thus, we have left to each of the contributors to the present volume the task of specifying the specific version of naturalism within his purview. Still, a cluster of positions and a set of philosophical commitments may fairly be associated with contemporary philosophical naturalism (hereafter, simply naturalism) as it has come to be widely understood by its friends and foes in the current philosophical setting.

Minimally, as a negative thesis naturalism implies that theism is false. Roughly, scientific naturalism is the view that the spatio-temporal universe of entities studied by the physical sciences is all there is. Positively, naturalism usually includes (1) different aspects of a naturalist epistemic attitude (for example, a rejection of so-called "first philosophy" along with an acceptance of either weak or strong scientism); (2) a Grand Story which amounts to an etiological account of how all entities whatsoever have come to be told in terms of an event causal story described in natural scientific terms with a central role given to the atomic theory of matter and evolutionary biology; and (3) a general ontology in which the only entities allowed are those that bear a relevant similarity to those thought to characterize a completed form of physics.

For most naturalists, the ordering of these three ingredients is important. Frequently, the naturalist epistemic attitude serves as justification for the naturalist etiol-

ogy which, in turn, helps to justify the naturalist's ontological commitment. Moreover, naturalism seems to require a coherence among the postulates of these three different areas of the naturalistic turn. For example, there should be a coherence among third-person scientific ways of knowing, a physical, evolutionary account of how our sensory and cognitive processes came to be, and an ontological analysis of those processes themselves. Any entities that are taken to exist should bear a relevant similarity to entities that characterize our best physical theories, their coming-to-be should be intelligible in light of the naturalist causal story, and they should be knowable by scientific means.

For our purposes, it is important to say a bit more about naturalist ontological commitments. A good place to start is with what Frank Jackson calls the location problem.[3] According to Jackson, given that naturalists are committed to a fairly widely accepted physical story about how things came to be and what they are, the location problem is the task of locating or finding a place for some entity (for example, semantic contents, mind, agency) in that story. Without passing judgment on Jackson's own analysis of the location problem, we maintain for at least two reasons that the naturalist ontology should, in some sense, be located in light of the naturalist etiology and epistemic attitude. First, if this is not done, then the naturalist ontology may well amount to a mere assertion and a controversial one at that, given the proliferation in recent years of robust alternatives to naturalism, e.g. theism. Second, the naturalist claims to have epistemic, explanatory, and methodological superiority on his or her side, especially when naturalist positions are understood to have the authority of science behind them. In order to justify this assertion of superiority, the naturalist ontology should be located in the naturalist etiology and epistemology.

Not all naturalists agree with this, though most would accept a location requirement of some sort, not altogether unlike the one mentioned above. More importantly, however, not all naturalists will agree regarding the resulting naturalist ontology that is so located. Here we may offer a pair of useful, if not exhaustive, distinctions regarding a naturalist ontology – global vs. local naturalism and weak vs. strong naturalism. Roughly, global naturalism is the view that the spatio-temporal universe of natural entities studied by science is all there is. Global naturalists (for example, Wilfrid Sellars, D. M. Armstrong) reject abstract objects of any kind, including sets, numbers, propositions, and properties/relations in general taken in the traditional sense as non-spatio-temporal abstract entities. Local naturalists (for example, Jeffrey Poland, Collin McGinn) either accept or are indifferent towards abstract objects but they insist that the spatio-temporal universe consists only of entities studied by the natural sciences. Local naturalists reject Cartesian souls, Aristotelian entelechies, and so forth.

We can also distinguish strong and weak naturalists. Strong naturalists (for example, David Papineau) accept strict physicalism for the natural world, while weak

naturalists (for example, John Searle) embrace various emergent entities. According to strong naturalists, naturalism constrains properties and relations just as much as it does particulars, and all entities whatsoever must be susceptible to exhaustive description in the language of an ideal physics (supplemented, perhaps, by chemistry and biology). What is clearly eschewed by strong naturalists is the need for psychological categories. Such categories may be useful for a number of purposes, but they do not carve the world at the joints. By contrast, weak naturalists have offered various forms of supervenience in order to cash out the proper requirements of naturalism. Whatever the form, the goal is to make supervenient entities dependent upon a subvenient physical base.

For someone to be a consistent naturalist of whatever form, he or she must hold that the entity which emerges or comes-to-be (as naturalists understand that notion) (1) fits solely within a naturalist ontology (that is, it bears a relevant resemblance to paradigm case natural entities of physics in its being and behavior), and (2) can, at least in principle, be given a naturalist etiological explanation in keeping with the naturalist epistemic attitude and Grand Story if it is to be explained at all. The naturalist ontology must be selected in such a way that it will cohere well with the naturalist epistemology and Grand Story in order for the naturalist to be justified in asserting the scientific explanatory superiority of naturalism over rival world-views.

In light of these considerations, critics and advocates of naturalism disagree about what ontology best expresses the naturalist world-view. Naturalist critic Reinhardt Grossmann has argued that the naturalist ontology should be taken to imply a rejection of abstract objects (such as sets, properties, propositions). Many naturalists accept this implication and either adopt some version of nominalism (for example, Hartry Field) or try to work out an acceptable form of naturalism that includes a realist construal of properties without taking them as abstract objects (for example, D. M. Armstrong). Likewise, many naturalists (for example, David Papineau) take naturalism most reasonably to require strict physicalism (all individuals, events, properties, relations, and so on are completely physical entities), though others (for example, John Searle) accept the existence of irreducible, genuinely mental, emergent or supervenient properties or events. Virtually all naturalists (for example, Daniel Dennett, John Bishop) eschew libertarian agency though naturalist William Rowe and anti-naturalist Tim O'Connor believe that emergent libertarian agents are consistent with an acceptable form of naturalism. The vast majority of naturalists eschew literal, irreducible teleology and substance dualism. Naturalists also have difficulty allowing for the existence of a number of features of the moral point of view, and they employ various strategies to handle problems in value theory.

All the contributors to this volume reject naturalism. Accordingly, we have a threefold purpose for the essays that follow.

First, the volume contains arguments that, if successful, justify the claim that naturalism fails to deal adequately with a number of desiderata (for example, its fails to provide an adequate account of acts of knowing, and it similarly fails to offer an adequate ground for or analysis of the preconditions necessary for scientific realism).

Second, we want to show that to the degree that naturalism is a consistent paradigm and its ontology is "located" within the physicalist story, to that same degree it should be a form of global, strong naturalism, as Papineau and others have correctly seen. Therefore, it should include rejection of abstract objects, irreducible mental entities, libertarian agency, and first philosophy. Those versions of naturalism that countenance abstract objects and knowledge thereof, supervenient mental entities, emergent libertarian agents, and so forth do so at the price of appearing *ad hoc*, of giving up the claim that science and science alone is adequate to explain (or is, superior to non-naturalistic attempts to explain) everything in one's ontology, of weakening and trivializing what it means to "locate" these entities, in contrast with naturalist successes in locating paradigm case macro-entities such as solidity.

As an illustration of this last point, there is a serious debate about naturalist employments of supervenience on at least two fronts: 1) Some question whether the supervenience relation itself (or states of affairs that contain supervenient entities) is an existent which a naturalist should countenance; (2) Others charge that supervenience is just a label for the problem naturalists must solve and not a solution.

It seems that naturalism faces a dilemma: either naturalism involves an epistemic attitude and etiology that express strong versions of scientism, in which case naturalism suffers from some obvious defects (no account of proper functioning, denial of consciousness) or else it must weaken its ontology to adopt certain entities (abstract objects, mental properties), in which case it loses the unity of science and its right to claim a strong naturalist epistemic attitude, explanatory hegemony, and an adequate etiological account of the coming-to-be of everything. This dilemma is part of the widely discussed explanatory gap or location problem for certain entities in the naturalist ontology.

Third, we want to provide a sampling of contemporary work in the burgeoning field of philosophy of religion which is pertinent to the debate over naturalism. In a recent retrospect of the twentieth century, Tyler Burge has remarked that "the central event" in philosophy during the last half century has been "the downfall of positivism and the re-opening of discussion of virtually all the traditional problems in philosophy."[4] Nowhere has this been more evident than in the field of philosophy of religion. Whereas philosophers of religion – what few of them there were – struggled during the fifties and sixties to defend the very meaningfulness of theological discourse in the face of positivist critiques, today's philosophical theologians find themselves in the midst of a renaissance of interest in religious perspectives on epistemol-

ogy, metaphysics, and science. All of the traditional arguments of natural theology, such as the ontological, cosmological, teleological and axiological arguments for God's existence, are matters of renewed discussion, along with theistic approaches to typical problems of philosophy normally pursued with a naturalistic methodology. We include in the present volume articles which advert to evidentiary considerations in the natural world itself, for example, contemporary Big Bang cosmology, indications of intelligent design discernible in the bio-world, and the apprehension of value through moral experience, which serve, as it were, as "signposts of transcendence" pointing beyond the natural world to its transcendent ground.

Finally, we wish to express our indebtedness to Lisa Vasquez, Nate King, Mercy Hung, Mike Austin, Richard Graziano, and Robert Garcia for their tireless efforts in preparing the typescript for publication.

Notes

1 David Papineau, *Philosophical Naturalism* (Oxford: Blackwell Publishers, 1993), 1.

2 Jaegwon Kim, "Mental Causation and Two Conceptions of Mental Properties," unpublished paper delivered at the American Philosophical Association Eastern Division Meeting, Atlanta, Georgia, December 27–30, 1993, pp. 22–23.

3 Frank Jackson, *From Metaphysics to Ethics* (Oxford: Clarendon Press, 1998), 1–5.

4 Tyler Burge, "Philosophy of Language and Mind," *Philosophical Review* 101 (1992): 49.

Part 1

Epistemology

1 Farewell to philosophical naturalism[1]

Paul K. Moser and David Yandell

Many philosophers have held, in the tradition of Plato and Aristotle, that there are uniquely philosophical, non-empirical methods of inquiry and that there are things whose investigation is reliably conducted via such methods. Rejection of any such methods or any such objects other than those available to sense experience or scientific methods traces back to the monism of the presocratic philosophers. Call any position entailing such a rejection "philosophical naturalism." (At times we will simply call such a position "naturalism," for short.)

Philosophical monism in general, including idealism as well as materialism, demands a *single* standard in metaphysics or epistemology, contrary to pluralism in metaphysics or epistemology. Explanation in philosophy and science is inherently unifying, subsuming a multiplicity of phenomena under classificatory unity – for example, under a unifying cause in the case of causal explanation. Explanation contributes unity and thus organization to what may otherwise appear as mere diversity. Proponents of monism, however, must attend to the risk of neglecting genuine data and truths resistant to a monistic explanatory scheme. What monism gains by unification of multiplicity in data may be lost by neglect of genuine recalcitrant data. Explanatory unity may be a virtue, but it will be virtuous only if pertinent truths and data are not excluded for the sake of theoretical simplicity. Since naturalism is, as we will see, monistic in its implications for metaphysics or epistemology, it faces the threat of neglecting genuine recalcitrant data. Contemporary motivation for naturalism derives largely from the success of the empirical sciences in representing their disciplinary phenomena as uniformly natural with respect to their being explained (in a sense of "natural" to be clarified below).

Philosophical dispute concerning naturalism ranges over many issues, but at bottom it concerns the nature of philosophy itself. A basic question is whether there is a legitimate form of philosophical procedure, often called "first philosophy," that has ontological authority but employs methods "prior to" or at least not based on sense experience or the empirical sciences. In particular, can a philosopher operating

without reliance on sensation or the empirical sciences legitimately engage in inquiry that posits real objects or at least yields genuine truths? Naturalists say no; antinaturalists, yes.

We will pursue not the history of disputes over first philosophy but rather the tenability of naturalism regarding first philosophy. In particular, we seek answers to the following questions: What motivates and supposedly warrants methodological and ontological naturalism? Does naturalism commit one to methodological or ontic presuppositions inconsistent with naturalism? If the methodology and ontology of naturalism aren't grounded in traditional first philosophy, can they be grounded at all, and if so, how? By attending to these questions and proposing a dilemma for naturalism, we'll contend that naturalism regarding first philosophy is unwarranted at best and self-refuting at worst. We will conclude, therefore, that naturalism is rationally unacceptable.

1. Naturalism: ontological and methodological

Ontological naturalism takes various forms. We will understand such naturalism in terms of this core view: *every real entity either consists of or is somehow ontically grounded in the objects countenanced by the hypothetically completed empirical sciences (that is, in the objects of a natural ontology).* As a general position about what actually exists, ontological naturalism is a metaphysical view, and not (just) an epistemological view about how we know or just a semantic view about the meaning of terms.

Ontological naturalists typically, *but not always*, endorse some kind of materialism (or, physicalism) about reality. So, two prominent construals of "material" are noteworthy. First, something is material (or, physical), according to some philosophers, if and only if it is extended in space. This approach portrays materialism as the view that everything that actually exists is extended in space, and thus that nothing non-spatial exists. This portrayal of materialism seems straightforward, but the relevant notion of "spatial extension" evidently depends on the very notion of "material" in need of elucidation. If there is such dependence, conceptual circularity will trouble the proposed characterization of materialism.

The apparent problem is that the notion of spatial extension is actually the notion of something being extended in *physical* space, or being *physically* extended. It is coherently conceivable that something has *temporal* extension, in virtue of extending over time, even though that thing is not extended in physical space. It is not self-contradictory, in other words, to hold that something is temporally extended but is not a body. The proposed characterization of materialism should thus be qualified to talk of *physical* space or *physical* extension. In that case, however, the problem of conceptual circularity is transparent. Even if there is no circularity here, the notion of spatial extension seems too close to the notion of being material to offer significant

clarification. Minimally, we need an elucidation of spatial extension, if talk of such extension is to elucidate talk of what is material.

The second prominent construal of "material" invokes the empirical sciences, as follows: the language of the empirical sciences determines what it is to be material, or physical. A predicate (such as "is a quark") signifies a physical item, on this view, if and only if the empirical sciences rely on that predicate in their formulations. Three noteworthy problems arise. First, the empirical sciences do not have an exhaustive list of truths about the physical world and are thus now incomplete; nor is it clear that they will ever have such an exhaustive list, given the complexity of the physical world. Such incompleteness allows for predicates that the empirical sciences do not rely on, but that nonetheless pick out something in the physical world. It is thus unclear that the empirical sciences have a monopoly on predicates signifying physical items. Second, the empirical sciences use mathematical predicates that evidently do not signify physical items. The predicate "is a member of a set," for instance, does not signify a physical item. Third, invoking "the empirical sciences" raises the problem of specifying exactly when something is an empirical science. This is no small task, given (a) the actuality of such highly theoretical empirical sciences as cosmology and particle physics and (b) the inclusion of the social sciences as empirical sciences. It would not help now to portray an empirical science as a discipline that investigates, in a special way, only the *material*, or *physical*, world, because we are after elucidation of the relevant notion of 'material'.

We now need elucidation of what it is to be material, or physical. Is social psychology, for example, an empirical science? Social psychologists often use mentalistic vocabulary (for example, talk of beliefs, desires, intentions, fears, and hopes) that does not obviously signify material things. So, we need clarification of the sense in which social psychology is materialist, if it is an empirical science. In general, then, ontological naturalists who are materialists would do well to characterize an empirical science without relying on talk of what is material, or physical.[2]

Ontological naturalists who are materialists oppose ontological dualism (the view that there are two irreducible kinds of things that actually exist) as well as other forms of ontological pluralism. The most influential kind of ontological dualism comes from Descartes, who affirmed that there are mental substances (particularly, thinking individuals) as well as physical substances (particularly, matter), and that mental substances do not depend ontically on physical things.[3] Proponents of Cartesian dualism must specify the sense in which mental substances are ontically independent of physical things. Cartesians allow for relations of *causal* dependence between the mental and the material, but they deny that mental things are part of the material realm. Being a mental thing, according to Cartesians, is different in kind from being a material thing. If there's a coherently conceivable distinction between minds and material bodies, it follows that materialism, understood as entailing mind–body identity, is not conceptually, or analytically, true – that is, true just in virtue of the

meanings of the terms "mind" and "body."

Any argument from a coherently conceivable distinction between mind and material body to an actual ontic distinction between the two may face trouble. An actual ontic distinction between minds and material bodies entails that actual minds aren't material bodies. This goes beyond a distinction between concepts or definitions; it is a difference between *things*. Even if it is not conceptually true that minds and material bodies are identical, according to many philosophers, it may still be contingently true that minds are material bodies. An identity relation between minds and material bodies may hold as an actual matter of fact, according to these philosophers, even if it does not obtain as a matter of meaning, definition, or necessity. Other philosophers deny that the notion of contingent identity is coherent and accept an inference from the coherent conceivability of mind–body distinction to the real distinction of mind and body. We need not endorse either position here. According to most contemporary materialists, the coherent conceivability of the falsity of materialism does not challenge the actual truth of that position. The *contingent* truth of materialism is adequate for most materialists.

Ontological naturalists who are materialists reject Descartes's ontological dualism, in particular, its implication that a human mind is composed of an immaterial substance different in kind from material bodies. They also reject Descartes's view that some mental properties are actually exemplified by certain *immaterial things*, things not spatially extended. Materialists are uniformly monistic in their view that all actually existing individuals are material, and in their view that only material things exemplify mental properties. These two fundamental materialist views, according to most contemporary materialists, are justifiable empirically, not a priori. The evidence for these views, in other words, must come from considerations dependent on experience. Neither reason nor definition alone can produce evidence for such views.

We should distinguish the following two views about the relation between ontological naturalism and materialism: (a) ontological naturalism is logically neutral regarding materialism, and thus is logically compatible with ontological dualism; (b) ontological naturalism logically presupposes materialism, and thus is logically incompatible with ontological dualism. Given (a), a defense of materialism on the basis of ontological naturalism should appeal to supporting evidence beyond naturalism, presumably empirical evidence. Given (b), an appeal to ontological naturalism in defense of materialism would be question-begging, as materialism would then be a logical prerequisite of ontological naturalism. Most contemporary ontological naturalists apparently prefer (a) to (b).

> W. V. Quine evidently would oppose (b) as follows:
> . . . nowadays the overwhelming purposes of the science game are technology and understanding. . . . The science game is not committed to the physical, whatever

> that means. . . . Even telepathy and clairvoyance are scientific options, however moribund. It would take some extraordinary evidence to enliven them, but, if that were to happen, then empiricism itself – the crowning norm . . . of naturalized epistemology – would go by the board. For remember that that norm, and naturalized epistemology itself, are integral to science, and science is fallible and corrigible.[4]

Quine thus suggests that naturalism is not logically committed to materialism even though naturalism does give primacy to fallible science. Quine's sense of "scientific" in this passage invites trouble. It plainly is not restricted to empirical science in the usual sense; for it implies that empiricism could "go by the board" as the result of our discovering evidence of parapsychological phenomena, but that this would not involve moving outside of "science."

A problem now emerges for Quinean naturalists. Quinean naturalists endorse a strong fallibilism, with good reason given the enormous shifts in the scientific conception of reality that have occurred within merely the last few generations. Just as a commitment to absolute space as the basis for material reality led many to oppose theories of space-time relativity, any a priori commitment about the nature of science could readily bring Quinean naturalists into conflict with the best contemporary science. Indeed, Quinean constraints on the ontology of the sciences are so loose that they permit the postulation of parapsychological processes, and Quinean constraints on methodology are so loose that they permit the jettisoning of empiricism. It seems, then, that the sciences are so unconstrained that they permit even non-empirical methodologies to ground belief in the immaterial. So, it is altogether unclear how non-empirical methods of the sort employed in first philosophy are to be ruled out at all by Quinean naturalism. After all, if what characterizes the sciences is the pursuit of understanding, and there is no requirement that the objects of understanding all fall within the range of sense-experience and scientific observation, then there is no result that scientific pursuit of understanding is to be constrained at all by empiricism.

The problem for Quinean naturalists stems from their reliance, for opposition to first philosophy, on a conception of science they deem "fallible and corrigible." In particular, they offer no reason to think that science, as merely a pursuit of understanding, should not include first philosophy. Indeed, given (1) that Quinean naturalists take any pursuit of understanding as science and (2) that at least one intelligible philosophical question (say, a question about the consistency of a position or about valid inference from a position) is not empirically resolvable but is nonetheless a pursuit of understanding, an inclusion of first philosophy by science seems required. The only way out for Quinean naturalists would be to deny (2), but given their unwillingness to commit themselves, as a matter of principle, to empiricism in epistemology, we see no ground for their denying (2).

Given the internal problem of Quinean naturalism, we will relate naturalism to *empirical* science, in keeping with the fact that naturalists do uniformly limit good scientific explanation to empirically relevant explanation. Without digressing to the complex task of supplying a precise account of empirical relevance, we will construe "empirical science" broadly to encompass any scientific discipline housed in a typical college of natural or social sciences. So, astronomy is in, and astrology out; psychology is in, and parapsychology out. Anthropology is in, but philosophy and theology are out, given their regular place in humanities or arts, owing to their being insufficiently empirical to be empirical sciences. Our criterion is rough, but it enables us to take steps nonetheless.

The previous core definition of ontological naturalism encompasses eliminative, reductive and non-reductive forms of ontological naturalism, which may be characterized as follows:

(i) *Eliminative ontological naturalism:* every real entity is capturable by the ontology of the hypothetically completed empirical sciences, and language independent of those sciences is eliminable from discourse without cognitive loss.
(ii) *Non-eliminative reductive ontological naturalism:* every real entity either is capturable by the ontology of the hypothetically completed empirical sciences or is reducible to something capturable by that ontology.
(iii) *Non-eliminative non-reductive ontological naturalism:* some real entities neither are capturable by the ontology of the hypothetically completed empirical sciences nor are reducible to anything capturable by that ontology, but all such entities supervene on entities capturable by that ontology.

Proponents of (i) have included W.V. Quine, Paul Churchland, and Daniel Dennett.[5] J.J.C. Smart, among others, has suggested (ii).[6] Donald Davidson and David Papineau, among many others, have suggested (iii).[7]

Views incompatible with positions (i)–(iii) are sometimes deemed naturalist. For example, some philosophers regard themselves as ontological naturalists even though they countenance the existence of non-physical abstract objects that neither consist of nor are somehow ontically grounded (by reduction or supervenience) in the objects of a natural ontology. They embrace ontological naturalism owing to their rejecting the existence of mental substances. Such a construal of ontological naturalism as a position in the philosophy of mind differs from the general ontological naturalism just identified. Nonetheless, many, if not most, philosophers rejecting Cartesian dualism in the philosophy of mind are also ontological naturalists in the core sense specified earlier.

Beyond ontology, the term "naturalism" picks out a range of views about the nature of legitimate *inquiry* that are logically independent of ontological naturalism. These views, falling under the general category of "methodological naturalism," share this

core position: *every legitimate method of acquiring knowledge consists of or is grounded in the hypothetically completed methods of the empirical sciences (that is, in natural methods).* This is not ontological naturalism, as claims about methods of acquiring knowledge are not the same as claims just about what exists. In agreement with most naturalists, we understand talk of the "methods of the empirical sciences" broadly, to encompass the kind of "lay scientific" observation underlying ordinary perceptual beliefs.

Methodological naturalism manifests a kind of variety analogous to that of ontological naturalism. We may thus distinguish:

(iv) *Eliminative methodological naturalism:* all terms, including empirically disputed terms (for example, normative and intentional terms), employed in legitimate methods of acquiring knowledge are replaceable, without cognitive loss, by terms employed in the hypothetically completed methods of the empirical sciences.

(v) *Non-eliminative reductive methodological naturalism:* all terms, including empirically disputed terms, employed in legitimate methods of acquiring knowledge either are replaceable, without cognitive loss, by terms employed in the hypothetically completed methods of the empirical sciences or are reducible (by, for example, definition, mutual entailment, or entailment) to terms employed in those methods.

(vi) *Non-eliminative non-reductive methodological naturalism:* some empirically disputed terms employed in legitimate methods of acquiring knowledge neither are replaceable by terms employed in the hypothetically completed methods of the empirical sciences nor are reducible to those terms, but the referents of these terms supervene upon those of the terms employed in the hypothetically completed methods of the empirical sciences.

W.V. Quine has suggested (iv), and Jerry Fodor has suggested a version of (vi).[8] Perhaps a reductionist like J.J.C. Smart would endorse (v).[9] Talk of "empirically disputed" terms in (iv)–(vi) connotes terms deemed by some to be insufficiently empirical to figure in the empirical sciences. Normative terms and mentalistic terms are familiar instances of empirically disputed terms. In (iv)–(vi), the word "knowledge" may be taken in the broadest sense, to include knowledge proper (however analyzed), justified or warranted belief, and even evidence for a belief. In other words, methodological naturalists claim that the hypothetically completed methods of the empirical sciences define or otherwise ground the only sorts of inquiry that yield knowledge, justified or warranted beliefs, and even evidence for a belief.

We do not have a simple recipe to identify the methods of the empirical sciences. Perhaps the best we can do is to refer to those methods employed in our best physics, chemistry, biology, anthropology, sociology, and so on. The debate will turn, quite naturally, to what exactly is picked out by "and so on" or by talk of our "best" empirical

sciences. We take it for granted that the methods of the empirical sciences, whatever specifically they include, are designed to be reliable, that is, truth-conducive. We do not deny that the methods of the empirical sciences seek other cognitive virtues, such as fruitful explanation and predictive success. We simply acknowledge that those methods aim at least to be reliable, regardless of whether they actually succeed. The turbulent history of the empirical sciences illustrates that they do not always succeed. That limitation, however, does not preclude *intended* reliability in the methods of the empirical sciences. Scientific methods aim for truth, even when they fail. We thus assume that every scientific method is epistemically relevant in that it aims for reliable conclusions. We turn now to some problems, including a troublesome dilemma, for naturalism.

2. A dilemma

Our dilemma will bear on positions (i)–(vi) inasmuch at it bears on the aforementioned core statements of naturalism satisfied by those positions, namely:

> *Core ontological naturalism:* every real entity either consists of or is somehow ontically grounded in the objects countenanced by the hypothetically completed empirical sciences (that is, in the objects of a natural ontology).
> *Core methodological naturalism:* every legitimate method of acquiring knowledge consists of or is grounded in the hypothetically completed methods of the empirical sciences (that is, in natural methods).

These are core statements of ontological and methodological "scientism," for they exalt the empirical sciences as the criterion for metaphysical and epistemological genuineness. They entail ontological and methodological monism in that they acknowledge the empirical sciences as the single standard for genuine metaphysics and epistemology. These core positions thus promise remarkable explanatory unity in metaphysics and epistemology, but is their promise trustworthy? For brevity, let us call the conjunction of these two positions *Core Scientism*, while allowing for talk of its ontological component and its methodological component.

Core Scientism is not itself a thesis offered by any empirical science. In particular, neither its ontological component nor its methodological component is a thesis of an empirical science. Neither component is represented in the empirical scientific work of either physics, chemistry, biology, anthropology, psychology, or any other natural or social empirical science. As a result, no research fundable by the National Science Foundation, for instance, offers Core Scientism as a scientific thesis. In contrast, the National Endowment for the Humanities would fund work centered on Core

Scientism. In addition, Core Scientism proposes a universality of scope for the empirical sciences that they themselves eschew. Individual sciences are marked out by the particular ranges of empirical data they seek to explain. Similarly, empirical science as a whole is marked out by its attempt to explain all *empirical* data and thus by the range of all *empirical* data. Given this constraint on empirical science, it would be surprising indeed if the empirical sciences had anything to say about whether nonempirical categories, the investigation of which lies outside those sciences, are empty. Sweeping metaprinciples about the nature of legitimate inquiry, particularly metaprinciples bearing on non-empirical inquiry, are not the fruits of the empirical sciences; they rather issue from philosophy, good or bad. Accordingly, Core Scientism is not the kind of scientific thesis characteristic of the empirical sciences. The empirical sciences flourish, have flourished, and will flourish without commitment to Core Scientism. In addition, antiscientism is in no way antiscience or antiscientific.

Friends of Core Scientism will remind us that it invokes not the current empirical sciences but rather the hypothetically completed empirical sciences. They will thus be unmoved by the absence of Core Scientism from the theses of the current empirical sciences. Still, the problem persists, because we have no reason whatever to hold that Core Scientism is among the theses of the hypothetically completed empirical sciences. A general problem is that predictions about what the completed sciences include are notoriously risky and arguably unreliable. The revolutionary history of the sciences offers no firm basis for reasonable confidence in such predictions. In addition, nothing in the current empirical sciences makes it likely that the completed sciences would include Core Scientism as a thesis.

The problem with Core Scientism is its monopolistic posture, a posture common among *philosophical* claims. It makes a claim about *every* real entity and *every* legitimate method for acquiring knowledge. The empirical sciences, for better or worse, are not thus monopolistic; nor do we have any reason to think that they will become so. Neither individually nor collectively do they offer theses about *every* real entity or *every* legitimate epistemological method. The empirical sciences limit their theses to their proprietary domains, even if wayward scientists sometimes overextend themselves with claims about every real entity or every legitimate epistemological method. Support for this claim comes from the fact that the empirical sciences, individually and collectively, are *logically* neutral on such matters as the existence of God, the reliability of certain kinds of religious experience, the objectivity of moral value, and the reality of thinking *substances*. Each such science thus logically permits the existence of God, the reliability of certain kinds of religious experience, the objectivity of moral value, and the reality of thinking substances. We have no reason to suppose that the hypothetically completed empirical sciences differ from the actual empirical sciences in this respect. Naturalists, at any rate, have not shown otherwise; nor has anyone else.

Proponents of Core Scientism may grant that the empirical sciences are not

logically monopolistic regarding real entities or legitimate epistemological methods, and thus that Core Scientism is not a thesis of the empirical sciences. Still, they can insist that Core Scientism is epistemically *warranted* by the empirical sciences. (A 'thesis' of the empirical sciences is, in our language, logically entailed by the empirical sciences, whereas a claim warranted by the sciences need not be thus logically entailed.) This move would turn the debate to principles of epistemic warrant appropriate to the empirical sciences. Specifically, what principles of epistemic warrant combine with the (hypothetically completed) empirical sciences to justify Core Scientism? More relevantly, are any such principles of warrant required, logically or epistemically, by the (hypothetically completed) empirical sciences themselves? No such principles seem logically required, as the (hypothetically completed) empirical sciences logically permit that Core Scientism is not warranted. Whether such principles are epistemically required depends on the epistemic principles warranted by the (hypothetically completed) empirical sciences, and the latter matter remains unsettled. We have, at any rate, no salient evidence for thinking that the (hypothetically completed) empirical sciences warrant principles of epistemic warrant that justify Core Scientism. The burden for delivering such evidence is on naturalists and remains to be discharged.

Minimally, the empirical sciences rely on abductive epistemic principles that certify inferences to a best available explanation of pertinent phenomena. The empirical sciences are, after all, in the business of best explanation. Even so, their domains of pertinent phenomena to be explained are not, individually or collectively, monopolistic in the manner required by abductive warrant for Core Scientism. For example, those domains do not, individually or collectively, preclude every kind of religious experience suitable to abductively warranted theistic belief. So, the abductive epistemic principles accompanying the empirical sciences will fall short of warranting Core Scientism, given the monopolistic assumptions of the latter. More generally, any domain of evidence outside the scope of the (hypothetically completed) empirical sciences will raise potential problems for the abductive warrant of Core Scientism, given its monopolistic assumptions.

By way of reply, some philosophers may propose that Core Scientism is constitutive of the hypothetically completed empirical sciences *properly understood*. The proposal suggests that the proper understanding of empirical science involves Core Scientism. This would amount to a vindication of Core Scientism by semantical fiat: it is just part of the proper understanding of empirical science that Core Scientism is correct. This solution is, however, too easy to be correct. It offers nothing to block critics of Core Scientism from making an analogous move, whereby a position contrary to Core Scientism becomes part of the *proper understanding* of empirical science. For example, one might as well build a theistic design hypothesis into the "proper understanding" of empirical science. Two, then, can play the unconvincing game of semantical fiat. If one objects that a theistic design hypothesis is insufficiently empirical to figure in the

proper understanding of empirical science, the same objection is invited by the ontological and methodological components of Core Scientism. If the proper understanding of empirical science rests on semantical fiat alone, Core Scientism will run afoul of various available "proper understandings" of empirical science.

Naturalists sympathetic to either pragmatism or some kind of antirealism may try to rescue Core Scientism via rejection of our characterization of scientific and philosophical understanding as truth-seeking. Pragmatists insist that the purpose of inquiry is to produce (theoretically) "useful" beliefs rather than to uncover truths. In addition, antirealists claim that the purpose of inquiry is to produce consensus regarding certain claims or to pursue any of a number of other non-alethic goals. (Pragmatism and antirealism face decisive self-referential problems, but we will not digress here.) The relevant problem with pragmatism and antirealism emerges from their assumptions about "the" purpose of inquiry. We have no reason to suppose that there is such a thing as "the" purpose of inquiry or that truth is no goal of inquiry. Even if inquiry has non-alethic goals of the sort favored by pragmatists and antirealists, it can still have truth as a central goal. Like those engaged in the empirical sciences, we are concerned with the nature of legitimate inquiry *in those cases where* inquiry is directed to the discovery of truth. That different standards (whether naturalist or otherwise) would govern inquiry with alternative goals is no surprise, and it has no relevance to the dispute over naturalism offered as a view about the nature of legitimate truth-seeking inquiry. It *may* be "useful" (assuming that this term avoids alethic implications) not to posit immaterial objects or not to employ non-scientific methods, but, even if this were useful, this fact would be irrelevant to the current dispute over Core Scientism, as we have defined this position. So, we will not concern ourselves further with pragmatist or antirealist defenses of Core Scientism.

We can now construct a troublesome dilemma for Core Scientism. Either Core Scientism is not itself a thesis included in or warranted by the (hypothetically completed) empirical sciences or its warrant depends on a special reformulation of the "proper understanding" of empirical science. In the first case, Core Scientism is self-defeating; in the second case, it is merely philosophically harmless. In either case, however, it is philosophically ineffectual. The self-defeat of Core Scientism would result from its failing to be included in or approved by its own proposed single standard for methodological and ontological integrity: the (hypothetically completed) empirical sciences. By its own proffered standard, in that case, Core Scientism would suffer defeat. If, however, Core Scientism is to be rescued via semantical fiat, as a desideratum for the proper understanding of empirical science, it becomes harmlessly stipulative in nature, with no firm basis in what actually goes by the name "empirical science." In that case, Core Scientism loses its ontological and epistemological teeth, inasmuch as it becomes divorced from the ontological and epistemological successes of the actual empirical sciences. Given this dilemma, Core Scientism becomes philosophically innocuous. We will explore the fate of naturalism

further, in light of the dilemma facing Core Scientism.

3. The logical status of naturalism

If any version of naturalism is true, it is either true necessarily or true contingently. If naturalism is true necessarily, the necessity is a product of either (a) the form of the characteristic claims, (b) the definitions of the terms employed, (c) the relations between the concepts underlying those terms, or (d) the necessary nature of reality. In short, the necessary truth of naturalism entails that its characteristic claims are either syntactic tautologies, definitional tautologies, conceptual tautologies, or metaphysical necessities. If the characteristic claim of a given form of naturalism does not fall under any of these categories, then it is contingently true, if true at all. We will argue that none of the forms of methodological naturalism under discussion falls under any of the categories of necessity (a)–(d), and that therefore methodological naturalism is at best contingently true. One might hold that ontological naturalism is necessarily true, but since this view is not defended by any prominent naturalist, we do not bother with it.

If methodological naturalism's characteristic claims are contingent, then any warrant we have for these claims will presumably be empirical. Given the problems identified in the previous section for the empirical warrant of methodological naturalism, the case for claiming that Core Scientism, including its methodological component, is contingently true faces serious trouble. If, then, no version of methodological naturalism is necessarily true, the case for methodological naturalism being true at all will likewise be in trouble.

Syntactic tautology and naturalism

None of the characteristic claims of the species of naturalism (i)–(vi) seems to be a syntactic tautology; neither, as expected, does Core Scientism. The denials of (i)–(vi) and Core Scientism seem to be consistent in logical form. So, the necessitarian naturalist has the burden of showing that a formal contradiction arises from denying a particular variant of naturalism or Core Scientism. Given the difficulty of discharging this burden, no prominent contemporary naturalist has defended naturalism on this basis.

Definition and methodological naturalism

A historically popular defense of methodological naturalism claims that such

naturalism follows from definitional truths. This defense proceeds on the basis of one of these options: (1) the terms in the aforementioned principles (iv)–(vi) could be stipulated to have certain meanings; (2) the common or conventional usage of those terms could give them particular meanings, or (3) those terms could pick out concepts having certain relations that yield definitional truth (this possibility will be treated in the next subsection).

Regarding option (1), any of the characteristic claims of naturalism could be made a definitional tautology by mere stipulation. If, for example, "legitimate method of acquiring knowledge" is simply defined as "method of the empirical sciences," then eliminative methodological naturalism's defining claim is a definitionally necessary truth. This strategy is, however, readily disarmed. One can grant the stipulation, and then shift the terms of the debate away from the stipulated terms by asking whether there are any methods for acquiring knowledge other than by means of "legitimate" (stipulatively defined as "empirical scientific") methods of inquiry. One could also frustrate such a strategy by refusing to grant the stipulation. Mere insistence that we so stipulate the meanings of terms is question-begging against opponents of this sort of necessitarian naturalism (including not only antinaturalists but also those holding that naturalism is contingently true and even those holding that it is non-stipulatively necessarily true).

A naturalist might claim that, as a matter of lexicographical fact, the terms employed in some version of naturalism are so used that the characteristic claim of that version is definitionally necessarily true. Given this claim, there is some standard or convention of linguistic usage on which some version of naturalism is true solely in virtue of the way the terms expressing it are used. Whether this constitutes a legitimate form of necessity is an open question, given that such standards or conventions of usage would themselves be contingent. We will offer reasons for doubting that methodological naturalism can be justified in this way regardless of whether such justification would yield genuine necessity.

Whether there are conventions regarding linguistic usage that make a form of naturalism definitionally true is an empirical matter, and must therefore be evaluated on the basis of empirical evidence. Naturalists, however, do not typically claim just that people *do* regard the terms of scientific methods as (or as grounding) the terms of the only legitimate methods of acquiring knowledge. Otherwise, the truth or falsity of naturalism would be established by linguistic survey, and this project would likely produce the overwhelming result "undecided" if popular linguistic usage is our guide. Rather, naturalists typically claim that people *should* adopt their naturalism. The principles of naturalism, including those of Core Scientism, typically have a normative character that would be lost if they were mere reports concerning how people happen to use certain words. In addition, the necessity and indeed the truth or falsity of naturalism would then fall prey to variations in linguistic usage. Naturalism could be necessarily true (or simply true) at one time or in one region but not in another. This,

however, would be an undesirable characteristic in any scientific or metascientific principle. The sciences seek truth, not just temporal or regional consensus.

A common claim regarding the source of normativity for naturalism is that it stems not (just) from ordinary usage of terms, but from *scientists*' usage. In other words, it enjoys support from expert linguistic usage. The problem, however, is that scientific linguistic usage does not enunciate monopolistic claims regarding the nature of legitimate inquiry, as Core Scientism does – even if wayward scientists occasionally engage in exaggeration via such monopolistic claims. Given the non-monopolistic conceptions of the empirical sciences standardly employed by scientists themselves, it's understandable that the sciences do not offer monopolistic answers to philosophical metaquestions about the nature of inquiry. If a scientist endorses Core Scientism, he or she is doing so not *qua* empirical scientist but rather *qua* metaphysician or epistemologist, in short, *qua* philosopher.

Philosophers claiming that the methodological component of Core Scientism is definitionally necessarily true face a dilemma: either they (a) seek to ground the definitional claim in the actual sciences by claiming that empirical truths about how scientists use methodological terms somehow (despite appearances) warrant such a metaclaim and somehow make it normative, or (b) seek to ground the claim in hypothetical future science that has expanded to cover such claims, by claiming that the methods of such science rule out traditional first philosophy as unscientific. Neither of these possibilities is promising.

Narrowly conceived, the empirical sciences include no discipline that makes even descriptive monopolistic claims about the nature of scientific practice. More broadly conceived, to include "science studies" or sociology of science, the empirical sciences will have a much larger range of methods available but still will not make *normative* claims, at least not monopolistic normative claims, about the legitimacy of methods of inquiry. For such normative claims to be included within the scope of the products of scientific methods, those methods would have to be expanded to include methods that employ normative terms and underwrite monopolistic claims about inquiry. If naturalism permits such expansion, the door is thereby opened to the legitimacy of competing forms of inquiry using normative terms, including those of traditional first philosophy. (This avenue is not open to any version of eliminative methodological naturalism that rejects as unscientific the needed normative terms.)

Suppose, as seems likely, that non-eliminative methodological naturalists cannot provide a standard governing the use of terms, *itself grounded in scientific method*, that permits the use of normative terms in metaclaims favoring naturalism while excluding as illegitimate antinaturalist conceptions of inquiry using normative terms. In that case, there will be no way that a definitional necessity can ground the rejection of antinaturalist forms of inquiry. Moreover, the normative terms employed within the needed naturalist standard, and those permitted by it, will have to be reducible to, or to refer to objects supervening on objects referred to by, the non-normative terms of

scientific inquiry. Otherwise, the retention of these terms will violate the standards of reductive or non-reductive methodological naturalism. We find no basis in the empirical sciences for the kind of standard needed by methodological naturalists.

The prospects for methodological naturalism, and thus for Core Scientism, seem bleak now. The empirical sciences, collectively as well as individually, are not in the business of making monopolistic normative claims about the legitimacy of methods of inquiry. One might seek to expand the empirical sciences to include methods for making such monopolistic normative claims. In that case, however, one would have made meta-epistemology a branch of the empirical sciences. Under such an expansion, much of what is now included in first philosophy would become "empirical science," owing not to a revision in first philosophy but rather to a sweeping liberalization of the idea of empirical science. Antinaturalists can consistently welcome such sweeping conceptual revision, given that it poses no threat to first philosophy. Naturalists, in contrast, will have to concede that such revision removes the epistemological teeth from Core Scientism, as it then allows first philosophy to proceed apace.

Conceptual necessity and methodological naturalism

One might defend the methodological component of Core Scientism on the ground that some version of it is a necessary truth owing to relations among the concepts underlying its constitutive terms. This modal status, so the reasoning goes, might arise solely from the way we conceptualize a legitimate method of inquiry. On one proposal, accordingly, the category "legitimate method" is conceived of as being included within, if not equivalent to, the category "empirical scientific method." On another proposal to the same end, there is a conceiving-independent relation of inclusion or equivalence between these two concepts. That is, one might hold that an objective conceptual necessity applies to the characteristic claim of some version of methodological naturalism. If, in contrast, the relation between the concepts in question is merely the product of convention or of some other conceiving-dependent process, then the option under consideration is equivalent to the possibility of definitional necessity discussed in the previous subsection.

A noteworthy consideration is that conceptual items are among the things whose reality is often called into question by methodological naturalists, particularly those of Quine's persuasion. Such items are intentional, and the conceptual relations being considered are semantic. If semantic or intentional items are needed to account for the logical status of methodological naturalism, based on conceptual relations, then eliminative naturalism cannot be necessary or even true on this basis, since it characteristically rejects the cognitive significance of such items. If the semantic and intentional features of the relevant concepts are not conventional, then the terms

denoting them are not plausibly viewed as reducible to non-intentional, non-semantic terms. Or, at least, naturalists will then need to make a special case for their reducibility, and the prospects seem bleak in light of the history of attempts at such reduction. So, reductive methodological naturalism is a dubious candidate for conceptual necessity (except where this is merely another name for definitional necessity, discussed above).

Non-eliminative non-reductive methodological naturalism allows for the legitimacy of intentional and semantic terms, but only if the intentional entities or features picked out by those terms and the semantic relations between those features supervene on non-intentional entities or features and non-semantic relations.[10] The non-eliminative non-reductive methodological naturalist faces a dilemma here. Either (a) the range of terms whose referents supposedly supervene on the referents of terms of the empirical sciences is broad enough to include the sorts of intentional terms needed to express metaphilosophical theories about meaning and to allow for reference to non-conventional semantic relations between concepts or (b) it is not. If it is, then the same latitude that makes non-eliminative methodological naturalism a candidate for conceptual necessity will permit competing views (including many that express traditional conceptions of first philosophy) to be candidates for the same status. If, in contrast, the range of legitimate terms is not this broad, then the characteristic claim of non-eliminative methodological naturalism is excluded from the domain of conceptual necessity along with its naturalist cousins. Still, only non-eliminative methodological naturalism (among forms of methodological naturalism) is a serious candidate for conceptual necessity. The problem is that the broad range of intentional and semantic terms needed as legitimate for this to be a live option, according to the view's own standards of legitimacy, would also permit (as legitimately naturalistic candidates for conceptual necessity) competing standards of legitimate inquiry that are not naturalistic in any sense that threatens first philosophy.

Metaphysical necessity and methodological naturalism

Many proponents of traditional first philosophy have held that certain principles are, although not analytic truths or tautologies, nonetheless necessarily true, in virtue of the necessary nature of reality. These principles are said to be "metaphysically necessary." Let us call any such principle a *substantive necessary truth*. Ethical principles, metaphysical principles concerning explanatory and causal sufficiency, and epistemic principles regarding the nature of evidence have been proposed as having this status. Most naturalists deny that there are any such truths. So, it would be rare for a methodological naturalist to appeal to such truths on behalf of naturalism. Still, we will challenge this option.

Could the methodological component of Core Scientism be a substantive

necessary truth? The claim that it is implies that (a) there is a necessary nature, or essence, of every legitimate method of acquiring knowledge and (b) this essence makes the methodological component of Core Scientism a substantive necessary truth. Suppose, then, that the methodological component of Core Scientism is the correct description of such an essence and thus is a substantive necessary truth. In that case, for the methodological component of Core Scientism to be itself a thesis or warranted product of the empirical sciences, the empirical sciences must be capable of including or warranting a substantive necessary truth of the monopolistic sort represented by Core Scientism's methodological component. We have already noted, however, that the empirical sciences do not trade in or warrant monopolistic methodological claims, let alone substantively necessary monopolistic methodological claims. Even if the empirical sciences include or warrant substantively necessary truths (and we're not going to stick our necks out here), they don't include or warrant *monopolistic* substantively necessary truths. Likely claims to empirical necessity, deriving from linguistic reference to natural kinds, are irrelevant to our point concerning the substantive necessity of Core Scientism. Linguistically grounded empirical necessities, familiar from the Putnam–Kripke–Marcus theory of linguistic reference to natural kinds, do not add up to genuine metaphysical necessity, owing to the undeniable contingency of linguistic reference.

In sum, then, first philosophy has nothing to fear from claims to the necessity of methodological naturalism. Core Scientism, we have seen, either collapses by its own standard for methodological genuineness or opens the door to the legitimacy of first philosophy. Let us turn, in conclusion, to an application of this dilemma.

4. The dilemma applied

W. V. Quine has proposed an influential version of Core Scientism implying that there is no cognitively legitimate philosophy prior to, or independent of, the empirical sciences, and thus that one should not make philosophical claims exceeding the empirical sciences.[11] Epistemology, by Quine's lights, is but a chapter in the book of empirical psychology; so epistemology offers no encouragement to first philosophy. Quine's own version of Core Scientism, his epistemological naturalism, is itself, however, an instance of philosophy prior to the empirical sciences. To meet this objection, Quine must show that his epistemological naturalism is a thesis or at least a warranted product of the empirical sciences. Quine does not discharge this burden, because the empirical sciences, as we have noted, do not make sweeping, monopolistic claims about the status of epistemology – even if an individual scientist makes such claims on occasion, in a moment of epistemological chauvinism. This may be an empirical truth about the empirical sciences, but it is warranted nonetheless, and it holds for the empirical sciences generally. Quine's epistemological naturalism,

combined with his general scientism, is thus self-defeating; by its own standard, it is omitted from the domain of the cognitively legitimate. Quine and other naturalists should seek to avoid self-defeat inasmuch as the empirical sciences do and inasmuch as theoretical conflict is disadvantageous to unified scientific explanation.

One might try to defend Quine's epistemological naturalism by offering a notion of *science* broader than that underwritten by the empirical sciences as standardly characterized in light of a typical college science curriculum. Such a proposal would enable Quine's epistemological naturalism to be a *scientific* hypothesis in the newly proposed sense of "scientific." This move would, however, yield a dilemma for naturalists: either (a) there will be a priori constraints on what counts as science (as actual linguistic usage of "science" would not ground this extended notion) or (b) this broader notion will be altogether *ad hoc* with regard to its rationale. To serve the purposes of naturalism, the new notion of science must exclude traditional epistemology and first philosophy, while including epistemological naturalism, in a way that is not *ad hoc*. This is no small task. In addition, this strategy for escaping self-defeat demands a hitherto unexplicated notion of science, which is likewise no small order. Neither Quine nor any other naturalist has defended such a strategy; nor have naturalists otherwise resolved the problem of self-defeat or the dilemma facing Core Scientism. At a minimum, naturalists will want to avoid a return to first philosophy, but it is not clear how they reasonably can.

An adequate epistemology, whether traditional or non-traditional, relies on *epistemological principles* that specify, among other things, what conditions epistemic warrant consists in. Quine's naturalized epistemology likewise relies on epistemological principles, such as his "maxim of minimum mutilation": a principle of conservatism implying that it is good not to alter our antecedent theory more than necessary.[12] Quine holds that another epistemological principle, concerning the maximization of "simplicity," can override his principle of conservatism; but this will not affect our discussion. We will bracket cases where the two principles conflict.

Is Quine's epistemological principle of conservatism itself a thesis of any of the empirical sciences? Quine has not shown that it is; nor has anyone else. People engaged in the empirical sciences perhaps *use* this principle at times, but no empirical science endorses it or seeks to confirm it. Again, the empirical sciences, given their theoretical objectives, do not concern themselves with endorsing or assessing general epistemological principles, particularly principles that are monopolistic in the way Core Scientism is. Quine's preferred species of epistemological naturalism thus includes a principle barred by his own scientism. In this respect, Quine's version of epistemological naturalism is self-defeating. As a result, it fails to challenge traditional epistemology or first philosophy.

Quine's epistemological principles evidently function as *analytic* truths for him, contrary to his own assumptions. The following questions invite this point: In virtue of what is Quine's epistemological principle of conservatism true, or correct? In virtue of

what is it actually the case that warranted theory revision conforms to the maxim of minimum mutilation? What, in other words, makes Quine's epistemological principle true rather than false? These questions highlight the importance of analyticity for epistemology, including Quine's naturalized epistemology. They ask what *makes it the case* that *warranted* theory revision is constrained by the principle of conservatism.

Apparently, it is just part of Quine's usage of the term "epistemic warrant" that the principle of conservatism constrains epistemic warrant. If so, it is *analytically* true for Quine that the principle of conservatism constrains epistemic warrant. A traditional alternative to this truth's being analytic proposes that it is a factually substantive principle about the nature of warrant, knowable a priori. This would be in keeping with the fact that the empirical sciences do not discuss, let alone endorse, any such epistemological principle. The latter alternative of substantive apriority would, however, be altogether alien to Quine's naturalistic project, violating both his scientism and his opposition to first philosophy.

Quine suggests that his epistemological principles find support in actual scientific practice. He claims that these principles "are maxims by which science strives for vindication in future predictions".[13] This suggests that Quine's epistemological principles are *justified* by their role in actual scientific practice. Quine might propose, in addition, that his principle of conservatism is *true* in virtue of its role in actual scientific practice. Let us grant this proposal for the sake of argument, and ask: In virtue of what is it true that actual scientific practice constrains epistemically warranted theory revision? It appears simply to be part of Quine's usage of "epistemic warrant" that actual scientific practice constrains epistemic warrant. In that case, it is analytically true for Quine that scientific practice constrains epistemic warrant. Appeal to scientific practice thus does not free Quine from the importance of analyticity. Barring analyticity, we would have great difficulty answering questions of the sort just raised. Perhaps this explains why Quine has very little to say on such questions.

Quine has commented on this question: What is it that makes one complete physical theory true and another false? He answers: "I can only answer, with unhelpful realism, that it is the nature of the world. Immanent truth, à la Tarski, is the only truth I recognize."[14] Even if we accept Tarski's disquotational approach to truth, this question persists: What makes it the case, at least for Quine, that the standard of conservatism constrains epistemic warrant? Talk of "the nature of the world" is unhelpful now, since we want to know what in particular about "the nature of the world" makes it the case, at least for Quine, that the standard of conservatism constrains epistemic warrant. The most plausible answer invokes what is analytically true for Quine regarding "epistemic warrant" (and related terms).

Quine, we have seen, uses principles regarding the nature of epistemic warrant in a way that treats them as analytic. Such principles, we have argued, are not themselves warranted by the empirical sciences; rather, one needs to use such principles to appeal

to the empirical sciences for warrant. Given this, and given that such principles are precisely of the sort characterizing first philosophy, we can plausibly understand Quine to be using analytic principles regarding warrant as an implicit first philosophy. The lesson is thus that if the analyticity of such principles involves first philosophy, as it seemingly does, then even naturalists such as Quine are wedded to first philosophy.[15]

Our dilemma for Core Scientism reveals, in the end, that Core Scientism is either self-defeating or ineffectively stipulative. We can nonetheless join naturalists in holding that empirical scientific methods are typically reliable for pursuing empirical knowledge of natural objects. Endorsing the empirical sciences as typically reliable in their appropriate domains, however, does not require embracing scientism of any sort. The epistemological principles justifying appeal to empirical scientific methods cannot themselves be warranted by appeal to those methods. Given this lesson, we may say not only farewell to philosophical naturalism but also welcome to first philosophy, as a foundational discipline here to stay.

Notes

1 We thank Paul Abela, Linda Mainey and Keith Yandell for comments.

2 For other difficulties in giving an informative characterization of materialism, see Tim Crane and D.H. Mellor, "There is No Question of Physicalism," *Mind* 99 (1990): 185–206.

3 For the argument that Descartes was a dualist, and not a metaphysical "trialist" as many recent commentators have claimed, see David Yandell, "Did Descartes Abandon Dualism? The Nature of the Union of Mind and Body," *British Journal for the History of Philosophy* (forthcoming). For discussion of the claim that Descartes conceded that interactionist dualism involved a contradiction, see David Yandell, "What Descartes Really Told Elisabeth: Mind–Body Union as a Primitive Notion," *British Journal for the History of Philosophy* 5 (1997): 249–73.

4 W. V. Quine, *Pursuit of Truth* (Cambridge, Mass.: Harvard University Press, 1990), 20–1.

5 See W. V. Quine, "The Scope and Language of Science," in Quine, *The Ways of Paradox* (New York: Random House, 1966), 215–32; Paul Churchland, *Scientific Realism and the Plasticity of Mind* (Cambridge: Cambridge University Press, 1979); and Daniel Dennett, *The Intentional Stance* (Cambridge, Mass.: MIT Press, 1987).

6 J.J.C. Smart, "Materialism," *The Journal of Philosophy* 60 (1963): 651–62.

7 See Donald Davidson, "Mental Events," in L. Foster and J. Swanson (eds.) *Experience and Theory* (Amherst, Mass.: University of Massachusetts Press, 1970), 79–101; and David Papineau, *Philosophical Naturalism* (Oxford: Blackwell, 1993). Roy Wood Sellars seems to endorse a version of this view, resting on "levels of causation" rather than supervenience, in *The Philosophy of Physical Realism* (New York: Macmillan, 1932).

For representation of the most prominent variants of contemporary ontological naturalism, reductive and nonreductive, see Paul Moser and J.D. Trout (eds.), *Contemporary Materialism* (London: Routledge, 1995). For assessments of contemporary versions of nonreductive naturalism, including those of Davidson and Papineau, see Paul Moser, *Philosophy After Objectivity* (New York: Oxford University Press, 1993), chap. 5; Paul Moser, "Physicalism and Mental Causes: Contra Papineau," *Analysis* 56 (1996): 263–7; and Paul Moser and J.D.

Trout, "Physicalism, Supervenience, and Dependence," in Elias Savellos and Ümit Yalçin (eds.) *Supervenience: New Essays* (Cambridge: Cambridge University Press, 1995), 187–217.

8 See W. V. Quine, "The Scope and Language of Science," in *The Ways of Paradox* (New York: Random House, 1966), 215–32; and Jerry Fodor, "Special Sciences," *Synthese* 28 (1974): 77–115.

9 See J.J.C. Smart, "Materialism," *The Journal of Philosophy* 60 (1963): 651–62. Cf. Smart, *Philosophy and Scientific Realism* (London: Routledge, 1963), chap. 3.

10 For the complexities involved in the phenomenon of supervenience, see the essays in Elias Savellos and Ümit Yalçin, eds., *Supervenience: New Essays* (Cambridge: Cambridge University Press, 1995).

11 See Quine, "The Scope and Language," 222; and Quine, "Things and Their Place in Theories," in *Theories and Things* (Cambridge, Mass.: Harvard University Press, 1981), 1–23, especially p. 21, for an identification of the scientific and the cognitive domains.

12 See W. V. Quine, *Pursuit of Truth*, 15, 56.

13 Ibid., 15.

14 W. V. Quine, "Responses" in *Theories and Things*, pp. 173–86; the quotation is from p. 180.

15 For further support for this lesson, see Paul Moser and David Yandell, "Against Naturalizing Rationality," *ProtoSociology* 8 (1996): 81–96.

2 Knowledge and naturalism

Dallas Willard

There is an *objective* difference between one who has knowledge of something and one who does not. This is true in both the occurrent and the dispositional senses of "knowledge" and "knows." That is, whether or not *X* has knowledge on a certain point or about a certain matter – knows the English alphabet, for example, the narrative content of *War and Peace*, or the date of Robert Kennedy's assassination – is not a matter of how anyone, including *X* himself, may think or feel about *X* and his conscious or other states.

Knowledge with reference to specific matters is a condition which individual human beings are or are not in. Usually this condition is found in a social context. But if only one person existed it would still be possible for that person to be knowledgeable about some, at least, of the specific matters that concerned him – for example, about reliable sources of food and water in his environment.

Can this objective difference among human beings consist in properties and relations that fit within a naturalist ontology? A long tradition of ancient and modern philosophers from Plato, Descartes and Kant to T. H. Green, Edmund Husserl and Hilary Putnam has insisted that it cannot. Others – especially those in the nineteenth and twentieth centuries who insisted upon distinguishing naturalism from materialism – have held that it could be. But with the rise and development of the mind/brain identity thesis during the last half of the twentieth century, the *generous* naturalism (as we shall call it) of Dewey, Santayana, Sidney Hook and others has largely disappeared in favor of a *narrower* naturalism more commonly and more correctly called "physicalism" (the older "materialism"). For it, all distinctively human properties are reduced to strictly physical properties of the central nervous system of the human body or to these plus characteristics of the natural and social setting – or, to nothing at all.[1]

In this chapter I will try to explain why narrower Naturalism or unqualified Physicalism cannot find a place for knowledge, and specifically for three of its essential components: truth, logical relations and noetic unity. At this late date it is hard to say much that will be strictly new on these matters, but, apparently, there is much that

needs to be said again. What I shall say about truth and logic is practically identical with what Frege said more than a century ago, though I hold views significantly different from his on how truth and logic fit into the full context of knowing and knowledge.[2] What I shall say about noetic unity adds little to what has already been said by Kant, Lotze and Husserl.

Naturalism is . . . what?

Before the question about the narrower naturalism's ability to accommodate knowledge within its permissible categories can be raised, one has to come to some understanding of what *Naturalism* and *knowledge* are, and this is difficult to do without begging important questions.

In traditional philosophical terms, naturalism is a form of monism. It holds, in some order of interdependence, that reality, knowledge and method each are of only one basic kind. That is, there are not two radically different kinds of reality or knowledge or method. It is fundamentally opposed to pluralism, and most importantly to dualism as traditionally understood (Plato, Descartes, Kant).

In its modern forms, naturalism further specifies its monism by reference to the empirical or the sense perceptible. The one type of reality admitted by it is that of the sense-perceptible world and its constituents. All knowledge is, for it, reducible to (or in some manner continuous with) sense perception, and all inquiry essentially involves sense perception, directly or indirectly. Currently, "the sense perceptible" is de-emphasized in favor of "the scientific" – the organization of data around empirically underdetermined hypotheses. But this is understood to constitute *empirical* research and, hopefully, to yield *empirical* or *descriptive* knowledge.

This leads directly into the current versions of "naturalized epistemology," where the emphasis is entirely upon the human being as a strictly physical organism acquiring beliefs in its "natural" environment. The question as to whether naturalism can accommodate knowing and knowledge is then replaced by the question of whether the various *normative* issues that arise with reference to knowledge and belief formation in the course of human life and scientific endeavors can be replaced by mere *descriptions* of actual processes of belief formation. In the famous words of Quine: "Epistemology [after being integrated into science through psychology] still goes on, though in a new setting and a clarified status. Epistemology, or something like it, simply falls into place as a chapter of psychology and hence of natural science. It studies a natural phenomenon, viz., a physical human subject."[3]

Knowledge itself, then – and, more weakly, justified belief – is simply belief that is produced in a certain way, for example, in ways that are *reliable*, ways that tend to produce true beliefs in actual as well as counterfactual situations that are relevant

alternatives to the actual situation.[4] Quine and others hold that epistemology *can* be appropriately replaced by psychology. This would yield what is now widely referred to as a "naturalized epistemology." Hilary Putnam, Jaegwon Kim and others hold, by contrast, that the normative (non-descriptive) element cannot be eliminated from epistemology, and therefore that a naturalized epistemology is impossible.[5]

All of this is now very familiar. But there are problems: problems which indicate to me, as it has to others, that the issue of the naturalization of knowledge is misconceived when stated in this way. And first of all, if we are going to replace epistemology with psychology we will have to decide *which* psychology will do. In fact, there is no existing psychology for the Naturalistic viewpoint to turn to. There is only an idealization of *some* adequate theory of human behavior which, supposedly, might at some time be achieved. This will, of course, have to be a "scientific" psychology. But what would that amount to? Could it possibly be a psychology that is not essentially identified in terms of naturalism itself? Highly unlikely. But then we are running in a circle, specifying naturalism in terms of psychology, but then . . . a *naturalized* psychology.

At present there are many socially identifiable, institutionalized forms of "psychology," with associated professional organizations, funding sources and avenues of publication. Research universities in the United States rarely ever have only one academic department giving advance degrees in psychology. So, when the naturalized epistemologist speaks of replacing epistemology with psychology, which department will he go to?

And we might also have second thoughts about normativity being *the* issue for the naturalization of epistemology. It is, of course, *an* issue. It is hard to imagine that the only epistemological question to be raised with respect to belief is its occurrence – even occurrence under certain conditions, no matter how fine-tuned those conditions may be. We would still want to know if our belief is "true," "correct," "right." Putnam asks: "Why should we expend our mental energy in convincing ourselves that we aren't thinkers, that our thoughts aren't really *about* anything, noumenal *or* phenomenal, that there is *no* sense in which any thought is *right* or *wrong* (including the thought that no thought is right or wrong) beyond being the verdict of the moment [or of some longer run], and so on? This is a self-refuting enterprise if there ever was one!"[6]

But let us surrender the point for a moment and grant that epistemology can be replaced by psychology, the normative by the descriptive. Will it make any difference with respect to the longstanding issues around the naturalization of knowledge? Not really. The descriptive/normative contrast has no essential bearing upon the issue of naturalism in epistemology. Getting rid of normativity will not secure the naturalization of epistemology and retaining it will not exclude naturalization. It has seemed to do so only because an important philosopher (Quine) in an influential

paper set that contrast up as the essential issue, and others, for whatever reasons, accepted his formulation as the framework of the debate.

On the one hand, a hardy naturalist such as Dewey might very well insist that there is nothing non-natural about normativity.[7] To think so is only to admit another of what Dewey often called "untenable dualisms." Norms are interwoven throughout human experience and are every bit as "natural" as anything else to be found there. And a Wittgensteinian reading of linguistic rules and forms of life could come up with much the same point. Only someone who reads "descriptive" (as opposed to normative) in a strongly materialistic or physicalistic fashion would take the opposition of normative to descriptive to be the issue of naturalism in epistemology. But many Naturalists have not taken "descriptive" in any such sense, and certainly science or experimental method does not of itself require them to do so. What, exactly, is so unnatural or non-natural about norms?

On the other hand, let us suppose we surrender norms – just banish "right" in both its adverbial and its noun forms from our treatment of knowledge and justified belief. What is so "naturalistic" about mere description? A description will be naturalistic only if what it is *about* – its *content*, what it mentions – fits into a naturalistic ontology. The issue for naturalism certainly is not *just* about normativity. What, for example, would make one think that an adequate description of the formation of belief by means of a reliable process would mention only things that would fit into a naturalistic ontology? Only, I suggest, the prior assumption that the "science" of psychology will, when fully perfected, be appropriately natural – that it will "emerge" from biology as biology from chemistry and chemistry from physics, or something like that. Or, at least, that it will mention nothing "dualistic" or "transcendental" with regard to the physical world.

And this does not even raise the issue of whether or not description as such can be understood in naturalistic terms, or whether describing itself is an entirely natural event or fact. A fully naturalized semantics is, once again, just presupposed in the move to "description" without norms. But would not the description (of reliable processes of belief formation, etc.) itself have to be "correct," "right," "adequate," "justified"? And have we then gotten rid of norms if we simply do not mention them in our descriptions of epistemic processes?

I presume that the age-old discussions in ethical theory pitting intuitionism and naturalism against each other (coming into the late twentieth century through the ethical intuitionism of G. E. Moore and developing in the subsequent critique of "descriptivism" by ethical theorists practicing "linguistic analysis") must have influenced Quine to frame the issue of naturalizing epistemology in the way he did. But his was a clearly unfortunate choice so far as genuine philosophical progress is concerned.

Invoking "science"

What has just been said with reference to psychology and naturalized epistemology calls attention to broader issues bearing upon the specification of naturalism.

Methodological monism is an enduring aspect of generic naturalism, and modern naturalism is often specified simply in terms of an exclusive application of scientific method in all inquiries. But how can this method support claims about the nature of reality as whole? For example, one might state that the only realities are atoms (quarks, strings, etc.) and derivatives thereof. But how is he to support his claim? It certainly cannot be derived from any specific science (physics, chemistry on up) or from any conjunction of specific sciences. And it is not to be derived through any application of experimental techniques within any science.

The naturalist must then have recourse to that popular but philosophically suspect abstraction, "science" itself, which says even less than the individual sciences about the nature of reality as a whole, because it says nothing at all. It is not the *kind* of thing that can say anything, though many individuals – usually, I think, not themselves scientists, and certainly not scientists expressing truths within the competence of their profession – present themselves as speaking for science, and thus as being "scientific" in some extended but still authoritative sense.

John Searle seems to be in this position. He speaks of "our scientific view of the world," which, according to him, every informed person with her wits about her now believes to be true. He speaks of a *view of the world* which includes "all of our generally accepted theories about what sort of place the universe is and how it works."[8] "It includes," he continues, "theories ranging from quantum mechanics and relativity theory to the plate tectonic theory of geology and the DNA theory of hereditary transmission," etc. We might imagine a very long conjunctive sentence – containing the specific theories he has in mind as conjuncts – that would, supposedly, express the "world view" in question.

But this will hardly do what he wants. One thing that will not show up in such a conjunctive sentence is *any* claim about reality as a whole or knowledge in general. Such specific scientific theories as those just mentioned – and no matter how many of them we may list – cannot provide an ontology. They never even attempt to determine what it is to exist or what existence is, and cannot by the nature of their content provide an exhaustive list of what ultimate sorts of things there are. Their existential claims are always restricted to specific types of entities as indicated in their basic concepts.

We emphasize the point that to suppose that a given scientific theory or conjunction of such theories provides an ontology constitutes a *logical* mistake, a misreading of what the theories say and imply. Those theories, and the bodies of knowledge wherein they are situated, actually say nothing whatsoever about *the*

universe or about how *it* – the whole "thing" – works. This is a merely semantical point about the meaning or logical content of the claims or sentences that make up the sciences. It is to be established or refuted by examining, precisely, those claims and sentences. It turns out that they do not even *mention* the universe, the totality of all that exists, nor do they say anything about the boundaries of knowledge generally. Such matters simply do not fall within the purview of their methods or findings.

In support of this claim we ask: Could one possibly find the place in some comprehensive and duly accredited scientific text or treatment, or some technical paper, where it is demonstrated or necessarily assumed by the science concerned that *all that exists* consists of particles or fields or strings – or whatever the proper subject matter of the science is? Would Searle or anyone else be able to mention the name of the physicist who established this as an "obvious fact of physics"?[9] Exactly where in the "atomic theory of matter" is the claim about what "the universe consists entirely of" to be found?

"After all," Searle rhetorically asks, "do we not *know from the discoveries of science* [italics added] that there is really nothing in the universe but physical particles and fields of forces acting on physical particles?" The answer, contrary to his assumption, is "No, we do not." Again, could he possibly just *point out* when, where, how and by whom this "discovery of science" was made. Was it made?

Also, before the philosopher can use "the discoveries of science" he must determine what "science" says. But this is to reify science, to treat it as an entity that issues "results." Science, as already indicated, says nothing at all. Particular scientists do. Unfortunately they also make unscientific statements. How can we tell when an individual scientist is making scientific statements, and "science" is therefore speaking, and when they are not? And can a "scientific" statement be false or perhaps illicitly derived and still be scientific?

If a scientific statement can be false or based on logical errors, then a scientific statement may be less than knowledge. How, then, could it be required that we accept such statements as a basis or framework for philosophical work? History shows that statements accepted as "scientific" have been both false and based on logical errors. Is the advocate of naturalism then one who works under an authority that may be and has been wrong? He himself would rarely if ever have the competence to do the scientific work and therefore must be taking the statements of "science" on authority. But blind authority is in fact one of the things we would expect naturalism to stand against. Historically it has done so, and that has been one of its virtues. How can it avoid resting on it, however, if what Searle says is true? And is a philosopher's statement about science, a scientific theory or scientist to be automatically regarded as itself scientific? What can *its* status be?

The dilemma of naturalism

Naturalism staggers back and forth between physicalism (materialism) as a general ontology and first philosophy, and outright physics-ism or scientism (which need not take the form of physics-ism) – often, though not always, trying to derive physics-ism from scientism and then physicalism from physics-ism. This continues up to the present.

In a recent review, Patricia Kitcher chides Stephen Stich for "philosophical Puritanism" when he takes naturalism to hold that the only real entities are physical.[10] Such a position apparently has now led Stich to give up naturalism "in favor of an open-ended pluralism." Pluralism, as he takes it, is a position that counts as legitimate all properties "invoked in successful scientific theories." But for Kitcher, it seems, such "Pluralism," tied to "successful science," is just the naturalism we want. She points out how "the obvious authorities" on naturalistic epistemology (Quine, Goldman) counsel us to "make free use of empirical psychology" and to "reunite epistemology with psychology."[11] Forget physicalism, her point seems to be. A loose scientism is enough to secure Naturalism for us. Indeed, many of the "generous" naturalists of the mid-twentieth century gathered around Dewey and Sidney Hook (see note 1) identified naturalism precisely with acceptance of science and only science as the arbiter of truth and reality, and seemed, at least, to accept whatever came out of the pipe of "scientific inquiry" as knowledge and reality.

But if the points made above about science, even "successful science," and about psychology in particular, are true, Kitcher's advice – similar to the advice of a Dewey or Hook – simply cannot be followed. It is vacuous in practice, for there is no way of identifying and accessing the "successful science" which is proposed as defining naturalism. At most you get "science now," which is really only "some scientist(s) now." And certainly no science (including psychology) that was not naturalistic in some strongly physicalistic or at least empiricist sense would be accepted as "successful" by those inclined to naturalism. Then we are back in the circle: naturalism in terms of science – but, of course, naturalistic science.

For these reasons, I take it that the appeal to science cannot serve to specify naturalism. There are, then, good reasons to be a "puritan" if you want to advocate naturalism. Naturalism has to be an honest metaphysics; and that metaphysics has to be "unqualified physicalism" as referred to above. But then a thinker who would be naturalist would feel pressure to have recourse to some specific a priori analyses to render his ontological specification of naturalism plausible. Short of that, one simply can find no reason why naturalistic monism with respect to reality, knowledge or method should be true: no reason why there should not be radically different kinds of realities with correspondingly radically different kinds of knowledge and inquiry. Why should sciences be "unified"? This lack of reason is, I think, what made A. E. Murphy

conclude long ago, in his review of *Naturalism and the Human Spirit*, "that the naturalists, who have so much that is good to offer, still lack and need a philosophy . . ."[12]

In addition to the difficulty of coming up with such a priori analyses, however, to turn to such inquiry as might produce them would be to break with the epistemological monism essential to naturalism and introduce something like a "first philosophy." This would be discontinuous with the empirical methods of the sciences. In showing its right through a priori analysis, naturalism would simply give up the game.

In specifying what naturalism is, therefore, one seems to be faced with an inescapable dilemma. Either one must turn to a priori (non-empirical) analyses to establish its monism, which will refute naturalism's basic claim about knowledge and inquiry, or its claim will have to rest upon a vacuous appeal to "science."

That might seem to end the discussion about naturalism as a philosophical alternative. But there may be a way to keep it going. One could retreat to a mere *methodological* naturalism and say that scientific method is our only hope as human beings. Whether or not we can adequately specify naturalism or know it to be true, one might say, the "scientific method" must be exclusively followed for human well-being. Naturalism would then be a humane *proposal*, not a philosophical claim. The proposal would be *to assume in our inquiries* that only the physical (or the empirical) exists and to see if inquiry based upon that assumption is not more successful in promoting human ends than any other type of inquiry.[13] Our task here would then be to show that the methodological assumption proposed contradicts what knowledge itself is, and to insist that it therefore cannot be an adequate methodological assumption, since it will not allow us to understand knowledge itself.

And knowledge is . . . what?

But now what about knowledge? I take knowledge in the dispositional sense to be identical with the capacity *to represent a respective subject matter as it is, on an appropriate basis of thought and/or experience.* In the occurrent sense it consists in *actually* representing, at a point in time, the respective subject matter as it is, on an appropriate basis of thought and/or experience. This is not intended as an analysis or definition of knowledge, but as an initial description of cases which count as knowledge or knowing.

What constitutes an "appropriate basis" will vary from subject matter to subject matter, of course, as is generally acknowledged of the corresponding methods of inquiry. It is, no doubt, impossible to define "appropriate basis" in any perfectly general way, or even to specify perfectly general necessary and sufficient conditions

for having an appropriate basis. Certainly I will make no attempt to do so here. However, a few things may be said about the *necessary* conditions of knowledge and knowing, without intending to be definitional or even comprehensive. The challenge to the narrower naturalism will be to accommodate these necessary conditions. If it cannot do so, it must be false, though it *might* still offer itself as a heuristic principle of inquiry.

Clearly one necessary condition of knowledge, both in the occurrent and dispositional senses, is truth. This is a necessary component of representing anything as it is. But this necessary element in any sufficient condition of knowing rules out all known psychologistic or sociological analyses of knowledge, such as certainty, rational acceptability, warranted assertability, reliability of process, etc., all of which apparently *could* be satisfied in conjunction with representations and beliefs that are false. Could, that is, unless truth is simply *defined* in terms of such psychological or sociological conditions. (Of course certainty, warranted assertability and the like each have psychological and other interests in their own right.)

We should also note that logical relations will be essentially involved in any case of knowing. That perhaps should be expected because of their intimate association with truth. To know one must *think*, in the sense of actively exploring the logical interrelationships involved in and with the respective representations. This is merely to say that the subject matter in question must be reflectively and thoroughly conceptualized, and the logical relations between relevant propositions and experiences carefully examined. Knowledge, accordingly, does not simply happen to a person. It is not a passion.

Finally, in order to know, one must have a certain broad familiarity with the subject matter itself – have observed it, tested it, informed oneself, thought it through – in ways suited to the kind of subject matter in question. How and to what extent this can and must be done will be determined by the relevant process of conceptualization of the subject matter (and of course by the subject matter itself), which will go hand in hand with observation and experimentation. All of this will produce a certain unity of consciousness that is necessary for there to be knowledge. Knowledge does not come in discrete units. Noetic unity across a complex field of consciousness within one person is required.

The absence of any reference to *belief* in my statements on knowledge and knowing will be immediately noticed. The absence is intended, and though rare today in discussions of knowledge it is by no means unique in the history of the theory of knowledge.[14] Even such a resolute naturalist as Roy Wood Sellars specifies the nature of knowledge without reference to belief.[15]

Belief I understand to be some degree of readiness to act as if such and such (the content believed) were the case. Everyone concedes that one can believe where one does not know. But it is now widely assumed that you cannot know what you do not

believe. Hence the well-known analysis of knowledge as "justified, true belief." But this seems to me, as it has to numerous others, to be a mistake. Belief is, as Hume correctly held, a passion. It is something that *happens* to us. Thought, observation and testing, even knowledge itself, can be sources of belief, and indeed should be. But one may actually know (dispositionally, occurrently) without believing what one knows.

Whether or not one believes what one represents truly and has an appropriate basis for so representing, depends on factors that are irrelevant to truth, understanding and evidence. It depends, one might simply say, on how rational one is. Now I do not think that this point about belief in relation to knowledge is essential to the rest of this chapter, but I mention it to indicate that the absence of any reference to belief in my general description of knowledge is not an oversight. Belief is not, I think, a necessary component of knowledge, though one would like to believe that knowledge would have some influence upon belief, and no doubt it often does.

In addition, it seems to me that specification of knowledge in terms of belief is a harmfully tendentious characterization, favoring the naturalization of knowledge. This is because belief has an essential tie to action, and is therefore easily located in the natural world – say as a mere tendency of the physical organism to behave in certain ways. I suspect that it is the almost overwhelming empiricist – and in that sense naturalist – tendency of thought in our time that has created the general presumption that knowledge must be some kind of belief. Hence we must here at least question that presumption; and, I believe, when questioned it will not prove to be obvious or, finally, sustainable.

It also should be noticed that on the view here advanced one may know without knowing that one knows. Skeptical tendencies are often associated with the mistaken view that one has to know that one knows in order to know. And, in particular, one does not have to know that one actually has "an appropriate basis in thought and experience" in order to have one, and one does not have to know that one's representation is true in order to know. It has to *be* true, indeed, but we do not have to know that it is true. We "use" it or live it; we do not "mention" or focus upon it.

Knowing that my representation is true is quite a different condition from knowing the respective objective circumstance to be the case. When I know that the book in my briefcase is a copy of Kant's *Critique of Pure Reason*, for example, what I know is not that my representation of the book in my briefcase as being a copy of Kant's *Critique* is true – though of course I *might* know that too. What I know is something about that book, namely, that it is in my briefcase. And likewise, when I know that, I may not know that I have an "appropriate basis" for representing it as I do – though I *might* know that too. It is enough that it *is* true, and that I *do* I have an appropriate basis for representing it as I do.

Clearly, if I cannot know without knowing that I know, then I cannot know that I know without knowing that I know that I know. And so forth. This is a genuinely vicious

regress. But we often do in fact know things without knowing that we know, and without even considering whether we know or not. Many familiar cases could be cited.

Essential to knowledge and knowing are, then, at least truth and logical grounding. We know only if our representations in the given case are true and logically non-arbitrary. Essential to logical grounding are logical relations: especially the simple formal relations of implication, consistency and inconsistency. Noetic unity comes in because knowing and knowledge require a larger context of consciousness involving many interrelated states and acts and kinds of states and acts. Our question now becomes: Can truth, logical relations and noetic unity be understood in physicalistic – and in that sense "naturalistic" – terms? Can knowing and knowledge be accommodated within the categories of physicalism, the narrower or "puritanical" naturalism?

Truth as "matching up"

Truth lies at the heart of knowledge. Knowledge is a condition of the human being that involves truth, for it involves representing its subject matter as it is; and, as we shall also see, much of our knowledge is in the first instance knowledge *of* truths.

Truth is also a vital human need, and a major part of what makes knowledge valuable. It and its opposite, falsity, are solidly at home in the midst of ordinary life. To know what truth is and to be able to recognize it and its opposite are basic components of ordinary human competence. To find one's way about, to communicate, and to give and receive directions and commands often requires us to identify the truth values of thoughts, beliefs and statements. In functioning in normal human relationships, say in a family or on a job, one must be able to recognize truth in thought, belief and statement. All of us bear a primary ethical responsibility to make sure that how we are thinking and speaking of things is as they are, that is, that our thoughts and words are true. Our view of the nature of truth must be compatible with its actual role in real life. But what is truth?

We first come to know truth – and what truth is – in concrete cases of verification within our physical environment. An infant in its second year of life or earlier develops the ability to look for something and to recognize *it* – what they are looking for – when it is found. The child at that point is capable of sustaining a specific thought or representation of something and of sorting objects that come before it with respect to whether they are or are not what they are seeking. Close to the same time the "uh oh!" phenomena emerges. The child observes things as *not* being how they 'should' be or are expected to be and verbally expresses the felt incongruity or lack of "fit." Closely linked with these developments is the ability to think of something *as* being such and such, and the associated capacity to find something to be (or to not be) as it is thought

to be. This is *verification* as a human reality. It is a primary form of knowing in the occurrent sense.

Soon the child learns the utility of lying, or representing things as it knows they are not. At that point it is in position to become truthful or honest, to routinely represent things as it knows them to be. Interestingly, children never have to be taught to lie. At an early age they figure it out quite on their own from their understanding of how thoughts and words do and do not match up with what they are about. This "matching up" – primarily of thought, and well before language is at the child's disposal – stands clearly before the child and is an essential condition of then learning the use of the words "true" and "false."

There is, accordingly, nothing esoteric, mysterious or enigmatic about truth (or falsity) itself, though many *particular* truths have such properties and some may even be completely unknowable. It is enigmatic or mysterious only for those who have decided that what it is it cannot be, or who have adopted a theory of mind (or language) and world that *makes* it either impossible or inaccessible. When what our belief or statement is about is as we believe or state it to be, when our representation or idea "matches up" to its object in the familiar way already indicated by cases, our representation (belief, statement) is *true*. Truth itself is just this characteristic of "matching up." When it is absent our representation etc. is *false*.

Truth and falsity are objective *properties* of representations. They are objective ways in which propositions differ and resemble among themselves, just as colors (red, yellow, green) and sizes are objective ways in which apples and other things differ and resemble among themselves. To say that they are objective is to say that they do not depend upon what we may think or feel about them. No proposition etc. can be made true merely by believing or favoring it – by one person or a million.

Now since this is so, what we find truth (or falsity) to be in the abundant cases where we *can* compare beliefs or statements (or, more properly, the propositions they involve) to what they are about is exactly what truth is in the cases where we do not or cannot directly compare thought with its subject matter. This characteristic constancy is something that truth values share with any object that, like truth, is not produced of modified by cognitive or other attitudes toward it. For example, thoughts to the effect that a certain candidate won an election, that the earth goes around the sun, and that Milton in *Paradise Lost* really intended to glorify rebellion are cases where we cannot directly verify or experience the truth of the thoughts by comparing them with what they are about. This inability, however, is due to the nature of the particular subject matter in relation to our cognitive faculties, not to the nature of representations, truth and reality as such. The truth of a belief or statement is not *created* by verification, but *discovered* by it. Otherwise we could prevent a belief from being true by refusing to verify it. Even in the cases most difficult to verify, truth remains a "correspondence" or "matching up" of the general type we become acquainted with in the verified cases.

For a thought or statement to be true, then, is simply for its subject matter *to be as it is represented* or held to be. When we confirm that a hitherto unconfirmed belief or statement is true, we do not *create* the relation (correspondence) it actually has to what it is about, any more that we create the fit of a wrench to a bolt head by placing the wrench on the bolt head, or the fit of a door to a frame by setting the door in the frame. The wrench *fits* the bolt head (or does not) even if it is never *placed upon* the bolt head, and the door fits the frame (or does not) even if it is never placed within it. And, similarly, a representation that is true is true even if it is never verified – by direct comparison with its object or otherwise. Truth is not the same thing as verification or proof, nor is it dependent for its existence and nature upon verification, any more than the fit or "correspondence" of the wrench to the bolt head is or is dependent upon the juxtaposition of the wrench upon the bolt head.

Also, what truth is does not change with time or historical process. It is a certain property or relation-like structure, and as such it is not the kind of thing that can change, any more than grey and yellow or sister or brother can – which is a totally different matter from how we choose to use the words "grey," "yellow," "sister," and "brother." When philosophers of the last two centuries have suggested that truth – this relation-like structure of correspondence that we all become acquainted with in our early years – is "really" the logical coherence or practical utility of beliefs or statements, etc., their suggestion is no more worthy of serious consideration than would be a suggestion that yellow is really an odor or that being a sister is the same thing as being a seamstress. Those suggestions were in fact based on the assumption that we *cannot* compare beliefs and statements with what they are about – an assumption that is refuted by the fact that everyone constantly does it.

The naivety of "correspondence"

But is not this view, as commonsensible as it may seem, simply naive, simple minded? Has not what we have learned about mind, phenomena, language and culture in the last two centuries shown that we do not and cannot experience such "truth," can never directly *find* the fit (or lack of fit) between representations and what they are of or about?

To answer this question in any remotely adequate manner would require a critical interpretation of theories of knowledge since Locke. It is, above all, with Locke that we seriously and continuously began to develop theories of representation (first mental and then linguistic or "symbolic") according to which our representations actually make it impossible to see whether our thoughts "fit" what they are about because any attempt to do so would only yield another representation, and so on. This theory has been the foundation for many of the major "triumphs" over naturalism,

from Kant and Hegel to Putnam and Derrida, as well as the root of the failure of many attempts at a genuine realist (often confused with naturalistic) theory of knowledge. It is also what has turned truth itself into an enigma or impossibility for many thinkers, and has led in the last two centuries to the emergence of many well-known candidates for its office.[16]

Still, a few points might be usefully mentioned in response to the charge of naivety.

First of all, the anti-correspondence, representationalist theories which now fill up the recent philosophical past are far from coming together in an adequate account of the mind–world relation or lack thereof. It is not as if there were now available some solid insight grounding an alternative to the type of accessible correspondence described above. In fact there is no generally acceptable alternative to correspondence. There is a series of successively discredited theories from Locke to Hume, to Kant to Hegel (or Fichte) to positivism and phenomenalism in their various forms; and then "language" (the "new way of words") is substituted for way of "ideas" or "experience," and the old battles fought over again. This time about how words tie to the world, and the outcome being a lingo-centric predicament instead of a ego-centric predicament. One cannot easily suppose that there is a philosophically credible alternative to the correspondence theory of truth. We do not have "something better" on hand.

Second, it is a noteworthy historical fact that every significant philosopher up to Kant accepted correspondence or "matching up" as the correct account of truth, even in cases, such as Hume's, where it was inconsistent with their overall system. Of course that does not by itself prove that that account is correct. But it is a historical fact that calls for some explanation.

This fact is, I think, associated with another significant fact: that those philosophers never suspected that thought was somehow linguistic, and that the real problem of truth had to do with sentences and what we do with them. Sentences and utterances do not, in fact, correspond to what they are about, and whatever truth may be as a predicate of them or other linguistic items is certainly open for speculation. What counts as a sentence (etc.) is also obviously relative to a language, and which sentences are true or false will be likewise. No doubt there are many interesting questions to be pursued with reference to sentences, languages, and "truth" as a linguistic predicate. But to suppose that a pragmatic or disquotational or rational acceptability theory of truth, for example, has anything to do with truth as a child comes easily to recognize it and as it is present in the constant experience of adults is simply to mistake or try to substitute one thing for another – of course from deep and powerful philosophical motivations.

These would surely include motivations of an empiricist or naturalistic character. Sentences and utterances, at least, present themselves as public, sense perceptible objects – though the "rules" governing them do not. But a certain version of

contemporary formalism in logical theory actually attempts to pass off the visible shape and arrangement of written symbols as logical form.[17]

Third, there remains the fact that we do constantly experience the "matching up" of thought with a subject matter we do not make or maintain in existence. At least we are strongly impressed that we do when we are not in the philosophical armchair: so strongly impressed that only some very strong philosophical impetus to the contrary could shake us. I think it is this that explains the long dominance of the correspondence theory among philosophers.

There are two main philosophical interpretations of consciousness that undermine this commonsense impression and the long philosophical tradition. One – the "Midas touch" picture of consciousness, as I call it[18] – is the view that to take something as our 'object' automatically transforms it in some essential way (possibly even making it 'mental'). How, exactly, consciousness – or for that matter language, or culture – being what it is, could make a tree or block of ice what it is, or turn something that was not already a tree or block of ice into one, is truly hard to say. We actually know how trees etc. come about, and they are not made by consciousness. One also can safely say that the story about how consciousness supposedly does its transforming and productive work has never been satisfactorily told. The second interpretation plays off of the saying that one *cannot escape* consciousness – cannot, as it is often said, "step outside of one's mind." Certainly, to be conscious of anything one must be conscious. But it does not follow from this that one cannot compare a thought to what it is about and see whether it "matches up" or not. Only confusion could make one think it does – a confusion probably based upon the "Midas touch" picture of consciousness.

Fourth, those who reject the correspondence model of truth do still, its seems to me, accept it, or at least its essential point, in a certain important respect. In advancing their own theories of the relationship between mind (or language/culture) and world they do not seem to me to suggest in any way that the truth of those theories about that relationship is in any sense relative. They seem to me to be telling us how things are with that relationship regardless of how they or any other person or culture may or may not speak or think of it. They give us the essential truth and the necessary essence of that relationship. I believe this to be true of the extreme relativists such as Rorty and Derrida, but also of the more modest ones such as Putnam.

Thus, Putnam states his own view that "truth and rational acceptability – a claim's being right and someone's being in a position to make it – are relative to the sort of language we are using and the sort of context we are in."[19] But can we imagine him to be thinking that his view of truth and rational acceptability here expressed is relative to language and context? I believe he intends to tell us what is the case with regard to truth and rational acceptability *itself*, and what indeed cannot be otherwise. That we are always "within" a language, he is telling us, is an *essential* truth that makes

"metaphysical realism" a strict and eternal impossibility. And yet he surely is, precisely, a metaphysical realist about our being within a language. The idea that in his discussions of language (mind) and world he is merely reporting on how things go from within a particular language and context which happen to be his does not seriously bother him – as surely it should if his views of truth etc. are, shall we say, *true*. But that would turn his view into something that would be of no general philosophical interest. And he does take philosophical interest to be of great importance. That is why he wants to be a "realist" of some sort, if only an "internal" one. He cannot just join up with Rorty *et al.* and treat truth as whatever can be sustained in your context. Thus his claim that Quine allows himself a "transcendental standpoint" in a certain respect[20] actually applies to himself, if I am right, with regard to his claims about rationality, truth, reference and allied subjects. Here, he indeed does take the "God's eye view."

Now this is by no means peculiar to Putnam, who is certainly one of the very finest of contemporary thinkers. I think it is built into the essential function of thought and assertion to present things as they are *without* regard to their being thought of or spoken of. Thus we have to use special forms of thought and language to express how things appear to us or are conceptualized by us or our group. If I am right, the simple correspondence indicated above has had the influence it has had over philosophical tradition and plays the role it does throughout human life because that is what really lies at the juncture between mind and world where we actually live. I believe no sound reason has ever been given for thinking it does not.

Why truth cannot be "naturalized"

Suppose, then, that truth in the sense of the matching up of representation with subject matter lies at the heart of knowledge. Can it be captured within the categories of the narrower Naturalism? I believe it clearly cannot. The argument against it is an old and simple one that has been reworked in many forms in recent decades.[21]

Suppose that we have an acceptable list of *physical* properties and relations. We might take them from physical theory, as the properties and relations corresponding to the concepts of current physics: location, mass, momentum and so forth. (Who knows what the future or ultimate physics will look like?) Or, moved by the above doubts about what philosophy can soundly derive from the sciences, we could turn to the primary qualities of modern philosophy, and, for that matter, add on the secondary ones as well: color, odor, etc. I don't think we need, for present purposes, to be very scrupulous about the list. Let us agree that whatever goes on such a list will count as physical properties, and that narrow naturalism is the proposal to confine our inquiries and conclusions to whatever shows up on the list and combinations thereof.

The argument, then, is simply that no such property or combination of properties

constitutes a representation of anything, or qualifies their bearer as being *of* or *about* anything. The properties of those properties and combinations thereof are not the same as the properties of representations (ideas, thoughts, propositions, beliefs, statements). If this is correct, and if the narrower naturalism admits only these properties, then there are no representations in the world of the narrower naturalism. Truth then disappears from that world, because in it no subject matter is represented; and hence it can never happen that something "is as it is represented or thought to be." With truth, knowledge also disappears. The ontological structure of knowledge cannot be present in the world of narrower naturalism.

Note that my claim is that such physical properties never *constitute* a representation. I say nothing here about representation (mental qualities) not *emerging* from the physical properties of, say, the human brain. This is not because I think they may so emerge, although some form of interaction between them and the brain surely does happen. Rather, it is because I can only regard talk of the emergence of irreducibly mental properties from the brain or the central nervous system as mere property dualism *cum* apologies.[22]

Significantly, Hilary Putnam and Daniel Dennett, in defending their own views of representation, belief and the intentional, emphatically support the view that the physical is devoid of the representational. Putnam asks us to imagine an ant crawling around in a patch of sand in such a way as to trace out "a recognizable caricature of Winston Churchill."[23] "Do the lines thus produced *depict* or represent Churchill?" Putnam asks. He thinks most people would say it does not. The ant has simply traced some lines "that *we* can 'see as' a picture of Churchill." Putnam's view is that nothing (in the brain/mind or out) *in itself* is a representation of anything, but is of or about (depicts or denotes), something other than itself, if it does so, only because *we* "take it" as depicting or denoting that other thing.

Dennett holds a similar view, and specifically with reference to the states of human beings, such as beliefs, desires, etc. In the typically "naturalistic" mode he declares his "starting point to be the objective, materialistic, third person world of the physical sciences," and holds "that philosophy is allied with, and indeed continuous with, the physical sciences."[24] Like Putnam (who of course is not a naturalist), depicting, denoting, etc. is for Dennett only a matter of how we treat something. "The intentional stance" or "intentional strategy," as he calls it, "consists of treating the object whose behavior you want to predict as a rational agent with beliefs and desires and other mental stages exhibiting what Brentano and others call *intentionally*."[25] The existence of belief, etc. with its intentionality is to be confirmed only by the success of the intentional stance as a strategy for predicting behavior. And when a better predictive strategy comes along, all the mental clutter of "folk psychology" will go the way of phlogiston and witches. Of course if that happens they are not "really there" now.

For Dennett, as for Putnam, there is nothing in the brain or out that by its nature

represents something else. There are no "natural signs." There are only human events of "taking as." And these events of "taking as" also, it would seem, must themselves lack any natural capacity for representation (of what is taken as). Rather *they*, in their turn, can only be *taken as* representing what is taken as intentional states or systems of the human organism, or what is taken as pictures or symbols. They do not inherently represent them.

Surely there is something wrong here. If we are in a world where nothing is naturally representative of something else, and we see the lines traced by the ant as a picture of Winston Churchill, then our seeing also is not naturally of the lines, and of the lines *as* depicting Churchill. Either there is going to be at some point a "taking as" which does not itself represent anything (even what is "taken") – which certainly sounds like a self-contradiction and is at best unlike the instances of "taking" featured in Dennet's explanations – or there is going to be an infinite regress of "takings." This inclines one to say that unless there are some *natural* signs – things that refer or represent simply because of what they are – there will be no *signs* at all.[26] But natural signs are, precisely, impossible in the world of strict physicalism – and, for his own reasons, in Putnam's more generous world as well.

Logical relations

If narrow naturalism cannot provide for truth, it also cannot provide for logical relations. Yet these too are essential constituents of knowledge. It cannot provide for such relations because they are, precisely, relations with respect to the truth values of propositions. Here we need only consider simple cases such as the relation of contrariety. Two propositions are logical contraries if they can both be false but cannot both be true. For example, *Sue's dress is red* and *Sue's dress is blue*. If one of these propositions is true, the other must be false. They cannot both be true. But both can be false – if Sue's dress is white, for example. The relation of contradiction, by contrast, is one that requires two propositions related by it to have opposite truth values, whichever they may be.

These and other logical relations are, like truth itself, objective relations. They obtain or do not obtain between propositions regardless of what any individuals or groups may feel or think about them. Moreover, laws expressing the logical relationships and logical character of propositions have a different sense and character from any laws of physical or psychological fact. They are neither hypothetical nor inductive, and have no existential import for such facts. They remain valid whether or not any such facts obtain. This becomes clearer if one tries to deduce or prove them from physical, psychological or linguistic facts or laws. It is not so much that it is not, in fact, done, or that it cannot be done, as that one cannot even imagine

what it would be like to do it.[27] These are points which Frege and Husserl elaborated so effectively in the late nineteenth and early twentieth Centuries that they could hardly be raised for discussion until the philosophical turn from thought to language and culture was more or less completed in recent decades. But I think they are points which can be made to stand up independently of the correspondence account of truth; and, if so, they provide a refutation of naturalism independent of that account and resting simply on the nature of the laws of logic.

But the centrality of logical relations for knowledge does not just concern the essential involvement of logical relations with truth values. Comparatively speaking, there is very little that we know we know because we are able to directly examine the respective subject matter and verify the truth of our ideas about it. This little is, of course, profoundly important, both in allowing us to understand what truth is, as explained above, and in providing true premisses from which we may proceed to other known truths by following out logical relations. Finding truths by following out logical relations occurs to an almost fantastic extent by the application of mathematics to various domains – and especially now, when such application is hugely extended by means of computers. In these cases, it is knowledge of the truth of the derived propositions that allows us to know that the corresponding state of affairs obtains and that its constituents exist. Here logical relations plus truth of premisses allows us to know unexperienced (and even unexperiencable) existence, rather than the comparison of existence to thoughts about it allowing us to know truths.

Noetic unity

As logical relations presuppose truth, so noetic unity presupposes logical relations – and more, presupposes a pattern of simultaneous and successive awarenesses that intercommunicates across a wide range of mental states and acts and their objects. The noetic aspect of knowing and knowledge encompasses all of the types of mental states and acts and their ways of coming together that are involved in the individual coming to know or to be in a state of knowledge. Knowledge is something that must be *possessed*. It is not the same thing as theory. And the possession of knowledge is an incredibly complicated and messy business, even with such a relatively simple part of knowing as a mere inference.

Thus Kornblith is quite right and makes a crucial point in rejecting purely apsychological accounts of belief (and knowledge) formation.[28] These types of accounts simply overlook the noetic requirements of knowledge and lead to what she calls "The Arguments-on-Paper Thesis," which treats justification as merely a matter of the logical relationships between propositions. The various forms of anti-foundationalism draw most of their ammunition from legitimate noetic

considerations, and that is their strength: Quine's famous "web of belief," Popper's picture of inquiry as driving piles into mud and sand just deep enough to support the bridge, Norwood Hansen's theory-ladenness of perception, Horkheimer's distinction between "traditional" and "critical" theory, the currently famous "intertextuality," etc.

Because the noetic issues are routinely mishandled through ontological mistakes or confusions carried into discussions of them – usually from empiricist or naturalistic assumptions – those who emphasize the noetic (and *in a sense* the psychological or even the social) conditions of knowing and knowledge most often turn out to be anti-realists. And one may be taken for an anti-realist just for attempting to do justice to the noetic requirements. About a century ago Wilhelm Wundt accused Husserl of refuting *logical* psychologism in the first volume of his *Logical Investigations* and reverting to it in the second volume – merely because the second volume is devoted to issues that are essentially noetic. The opposite of naturalism, however, is not some form of subjectivism or idealism. Anti-naturalism does not need to downgrade the ontological status of the empirical world and the empirical self. We don't surrender realism to defeat naturalism.

Still, for all the risks, any careful practice of knowing or theoretical examination of knowledge (and none have surpassed Husserl's) must take its noetic dimensions into account. That is why philosophers do not write their arguments on three by five cards and pass them around to each other, and why even footnotes (*pace* Searle) are not a sign of low philosophical quality.[29] Philosophers produce expository texts, as well as arguments, to express, convey and evoke a noetic context within which actual understanding and knowing can occur and arguments be appreciated.

Thus, in his *Discourse on Method*, Part II, Descartes lays out his favored four from among "the great number of precepts of which Logic is composed." All four are clearly noetic principles, from refusing to "accept" whatever is unclear, to the meticulous review of the steps in the progression from the clear and simple to the complex but logically derivable. Of course they also presuppose truth and logical relations and an ordered awareness thereof.

In a more recent philosophical presentation of modern logic, L. S. Stebbing distinguishes and interrelates *inference* (the essential noetic structure) and *implication* (the logical relation), the "therefore" and the "if . . . then," as she also puts it.[30] (She uses the term "epistemic" but clearly means the noetic as here understood.) The "epistemic" conditions of coming to know q on the basis of p, as she describes them, include: p must be known to be true, and p must be known to imply q without its being known that q is true. She also holds that "although p may *imply* q when q is false, yet q cannot be validly inferred from p unless it is the case both that p is true and is known to be true." She goes on to discuss the difference between inferring something ("validly"?) and tracing out the logical consequences of a proposition not assumed to be true or even assumed to be false, as in *reductio ad absurdum*. Such use of assumptions

and logical relations are obviously a crucial part of coming to possess knowledge, though they clearly are not inferences or arguments.

Alvin Plantinga rightly notes that "foundationalism is a *normative* thesis about noetic structures."[31] And, of course, the same is true about the many versions of Coherentism, verificationism, *Aufbau* projects, social constructionism, externalism, internalism, and "linguistic rule" theories that have turned up during the last century or so.

So, can the noetic be naturalized? It clearly cannot be, because of its pervasive involvement with logical relations. But there is more. The Kantian theme of the "unity of apperception" turns up here. His primary focus was upon the unity of a simple judgment and what was thereby presupposed about the mind. That is an important topic in itself, but knowledge and knowing does not come in the form of a simple judgment, or two or three of them. It comes in the form of a vast "web," to borrow Quine's word, of judgments, conceptualizations, perceptions, memories, even feelings and sensations, and experiences of many kinds. Indeed, it is not too much to say, a web of *life*. To know, we have said, is to represent something as it is, on an appropriate basis of thought and experience. The "appropriate basis" is never just awareness of a few logical relations, and a considerable "background" must be in place before the simplest cases of verification – finding the broom to be in the closet as I thought it to be – can occur.

Now, without logical relations and awareness of logical relations none of this could exist. But everything from how rationality functions across the great noetic web of the individual self, to the basic nature of creative genius (what *was* it about Einstein, after all?), to issues of self-identity are involved in noetic unity. This, I believe, is why it is impossible to lay down general sufficient conditions of an "appropriate basis." And before naturalism can triumph it would have to provide an elucidation of noetic unity within the framework of physicalism specified above. This will take more than an argument that pain is a chemical process in the brain plus a salute to the future of brain science. We will need an "identity thesis" that reduces Einstein's understanding of physical theory to brain states before it becomes really interesting.

Summary

We have tried to show why naturalism must be taken in the form of a "puritanical" physicalism if it is to be a philosophically significant position, and have presented knowledge as involving at least truth as correspondence, logical relations and noetic unity. We have argued that there is no place for truth or logical relations in a world where the only properties are physical, and therefore that noetic unity is also impossible in such a world. Since it is possible – many things are known and there are

people of great knowledge – naturalism must be false. It cannot accommodate the ontological structure of knowing and knowledge.[32]

Notes

1 What might be called "generic naturalism" has a long history that includes: classical naturalism, with figures such as Democritus, Epicurus, Aristotle and Lucretius; Renaissance naturalism, with Bruno, Campanella and Telesio, and – born too late – Spinoza; empiricist/nominalist naturalism, with Hobbes, Hume, D'Holbach and most of the French Encyclopedists and Comte; nineteenth-century materialistic naturalism, with Jakob Moleschott, Karl Vogt, Ernst Haeckel, Ludwig Büchner, Herbert Spencer and, it is often presumed, Charles Darwin; mid-twentieth-century (largely anti-materialistic) naturalism, with Santayana, Dewey and others; and late-twentieth-century ("identity thesis") naturalism, which wavers between scientism and physicalism, with Quine, David Armstrong, Paul and Patricia Churchland, John Searle, etc.

To appreciate contemporary naturalism for what it is, and the logical nuances that surround it, one has to see it in this long historical context. The single unifying theme of all Naturalisms is *anti-transcendentalism.* Their steady point of reference is the visible world and whatever it contains, which is "nature" in extension. Nothing "outside" it is to be allowed. This visible world is held to be self-existent, self-explanatory, self-operating and self directing. Usually though not always it is thought to consist entirely of processes involving only blind force. But what "nature" is in intension has never been agreed upon among naturalists. Some look very much like pantheists, and yet others (Santyana, Dewey) reach very far to incorporate "the divine" and all that is humanly unique into "nature." (See, currently, the divergence between Searle and, e.g., Paul Churchland or Daniel Dennett on the nature of the the mental.) Thus "self" in "self-existent" etc. only has the negative meaning of "non-other," i.e., not in virtue of something separate from this thing called "nature."

Effective entry into the long story for use by a contemporary thinker can be gained by starting with the article in *Encyclopaedia Britannica*, 10th ed., s.v. "Naturalism," by James Ward, 88, and then going on to his *Naturalism and Agnosticism*, (London: A & C Black, 1915). W. R. Sorley's *The Ethics of Naturalism* (London: William Blackwood and Sons, 1904), especially 17–21, is also helpful in understanding how naturalism has tried to distance itself from materialistic naturalism of Vogt, Haeckel, Büchner, etc. A series of articles on naturalism in *The Journal of Philosophy* from 1945 through 1949, easily identifiable by their titles, was evoked by the appearance of *Naturalism and the Human Spirit* [ed. Yervant H. Krikorian (New York: Columbia University Press, 1944)] and by A. E. Murphey's excellent critical review of it in *The Journal of Philosophy* 42 (1945): 400–17. The outcome of the mid-twentieth-century discussion is nicely summarized by Arthur Danto's article, "Naturalism," in Paul Edwards, ed., *The Encyclopedia of Philosophy*, vol. 5 (New York: Macmillan Publishing Co., 1967), 448–50. One of the intriguing aspects of the current situation is how Materialism, which was thought to be dead or something to be avoided for the first half of the twentieth century, came to life again in association with the "identity thesis" of mind and body and a new scientism, and led to a reformulation and resurgence of naturalism at the end of the twentieth century. Reading

Danto's fine article you would never have thought it possible.

For broader, cultural bearings of naturalism, see Paul F. Boller, Jr., *American Thought in Transition: The Impact of Evolutionary Naturalism, 1865–1900* (Chicago: Rand McNally & Company, 1969), and John Ryder, ed., *American Philosophic Naturalism in the Twentieth Century* (Amherst, N.Y.: Prometheus Books, 1994).

2 See for details my paper, "The Integrity of the Mental Act: Husserlian Reflections on a Fregian Problem," in *Mind, Meaning and Mathematics*, ed. Leila Haaparanta (Dordrecht/Boston: Kluwer Academic Publishers, 1994), 198–224.

3 W. V. Quine, *Ontological Relativity and Other Essays*, (New York: Columbia University Press, 1969), 101.

4 Hilary Kornblith, "Beyond Foundationalism and the Coherence Theory," in Hilary Kornblith (ed.) *Naturalizing Epistemology*, 2nd ed., (Cambridge, Mass.: The MIT Press, 1994), 142.

5 See Jaegwon Kim, "What is 'Naturalized Epistemology'?" in, *Naturalizing Epistemology*, 33–55; and Hilary Putnam's two Howison lectures, published in his *Realism and Reason* (Cambridge: Cambridge University Press, 1983), Chapters 12 and 13, under the titles "Why There Isn't a Ready-Made World" and "Why Reason Can't Be Naturalized."

6 Putnam, *Realism and Reason*, 246.

7 See the chapter "The Construction of Good," in Dewey's *The Quest for Certainty* (New York: G. P. Putnam's Sons, 1960) or most any of Dewey's mature works, such as *Human Nature and Conduct* (New York: The Modern Library, 1922) or, perhaps best of all, *Experience and Nature* (New York: Dover Publications Inc., 1958). On norms in nature see also Alvin Plantinga, *Warrant: The Current Debate* (New York: Oxford University Press, 1993), 72, and the thorough discussion in his *Warrant and Proper Function* (New York: Oxford University Press, 1993), Chapter 11.

8 John Searle, *The Rediscovery of the Mind* (Cambridge, Mass.: The MIT Press, 1992), 85.

9 Searle, *Rediscovery*, xii.

10 Patricia Kitcher, "Review of Stephen P. Stich's *Deconstructing the Mind," The Journal of Philosophy* 95 (December, 1998): 641–4, 641–2. This entire review is actually quite instructive on the problems of specifying naturalism currently.

11 Ibid., 642.

12 A. E. Murphey, "Book Review of *Naturalism and the Human Spirit," The Journal of Philosophy* 42 (1945): 400–17, 17. This may be the deeper reason why there is now a widespread sense that, in the words of Michael Friedman, naturalism "has reached the end of its useful life." See his thorough examination of the current situation regarding naturalism in his Presidential Address to the American Philosophical Association: "Philosophical Naturalism," *Proceedings and Addresses of the American Philosophical Association* 71 (1997): 7–37, as well as Barry Stroud's Presidential Address of the previous year, "The Charm of Naturalism," *Proceedings and Addresses of the American Philosophical Association* 70 (1996): 43–55, which emphasizes the lack of anything approaching a consensus concerning the meaning of the term. The volume *Naturalism: A Critical Appraisal*, Steven Wagner and Richard Warner, eds., (Notre Dame, Ind.: University of Notre Dame Press, 1993) also demonstrates what hard days have now befallen the movement.

13 One can also retreat to agnosticism, as seems to have happened with Thomas Huxley and

others in the late 1900s. See the article, "Naturalism," in *Encyclopedia of Religion and Ethics*, ed., James Hastings (New York: Charles Schribner's Sons, n.d.), 195–8. Also James Ward, *Naturalism and Agnosticism*. Agnosticism about matter has begun to re-emerge in the philosophy of mind in recent years.

14 For a quick introduction see the article by Steven Luper-Foy, "Knowledge and Belief," in Johathan Dancy and Ernest Sosa (eds.) *A Companion to Epistemology* (Oxford: Blackwell Publishers, 1993), 234–7. Unfortunately, John Cook Wilson is not discussed there.

15 Roy Wood Sellars says that "knowledge is the possession of ideas which do *disclose* the characteristics of the object denoted. In knowing we hold ourselves to grasp the nature of the object, its properties, characteristics.… Knowledge is the disclosure of the characteristics of existence." Roy Wood Sellars, *The Philosophy of Physical Realism* (New York: The Macmillan Company, 1932), 106. The entire Chapter IX, "Knowing a Common World," should be read.

16 First Hegel and Kierkegaard, then pragmatism and "coherence", then several twentieth-century options parasitical upon language (verifiability, rational acceptability, disquotational theories, and so forth). Loosening the grip on truth as an objective structure quickly led to loss of an objective and imperious logic. See the ways Hume, Kant, Hegel and Mill abuse the idea of logic, and for later developments up to the present see my "The Degradation of Logical Form," *Axiomathes* (1997): 31–52.

17 See my "Space, Color, Sense Perception and the Epistemology of Logic," *The Monist*, 72 (1989): 117–33, for a critical discussion of this idea. With respect to the linguistic nature of thought, Dennett remarks: "So the argument for a language of thought comes down to this: What else could it be?" [a]Daniel C. Dennett, *The Intentional Stance* [Cambridge, Mass.: The MIT Press, 1987], 35 and see the footnote on this page). Well, there are some interesting possibilities as to what thought apart from language might be, well known in the history of thought, if one can but for a moment escape the grasp of empiricistic naturalism.

18 See my "Predication as Originary Violence: A Phenomenological Critique of Derrida's View of Intentionality," in *Working through Derrida*, ed., Gary B. Madison (Evanston Ill.: Northwestern University Press, 1993), 120–36.

19 Putnam, *Realism and Reason*, 234.

20 Ibid., 242.

21 In the last chapter of his *The Intentional Stance*, Dennett gives a very helpful presentation of the issues in recent discussions, from the discussions between Roderick Chisholm and Wilfrid Sellars to the time of the writing of his book in 1987.

22 Searle's position is the clearest case of this. See *The Rediscovery of Mind*, and the more recent collection of papers, *The Mystery of Consciousness* (New York: The New York Review of Books, 1997). I do believe that emergence can be employed as a valid and useful concept in numerous domains, e.g., chemistry, sociology and the arts. But its valid employment requires some degree of insight into why *this* emerges from *that*. Such insight is lacking in the case of the brain and thoughts. This is a basic point made by such authors as Thomas Nagel and Colin McGinn. Searle's "simple solution" to "the famous mind–body problem," which "in a sense, we all know . . . to be true" simply refuses to face up to this fact (*Rediscovery*, 1). Rightly insisting on the irreducibility of mental properties, Searle tries to force them to be natural by assigning them a role in evolutionary theory and claiming that they in some literal sense are present in or on the brain. As to the former point, even Descartes' could have recognized the

function of his mental qualities in survival. Of course he did not think they 'emerged' from the brain. But neither Searle or anyone else has given any sense to the claim that they come from or are in or upon the brain. The actual relationship remains totally obscure.

23 Hilary Putnam, *Reason, Truth and History* (Cambridge: Cambridge University Press, 1981), 1.

24 Dennett, *The Intentional Stance*, 5.

25 Ibid., 15.

26 This is the thesis developed in Laird Addis, *Natural Signs* (Philadelphia: Temple University Press, 1989), especially "Part Two."

27 For development of this point with special reference to Quine, see my "The Case against Quine's Case for Psychologism," in *Perspectives on Psychologism*, ed. M. A. Notturno (Leiden: E. J. Brill, 1989), 286–95. And see the lengthy discussion of psychologism and the laws of logic in my *Logic and the Objectivity of Knowledge* (Athens, Ohio: Ohio University Press, 1984), 143–66.

28 Hilary Kornblith, "Beyond Foundationalism and the Coherence Theory," 133. She seems unaware that she is dealing with issues that were thoroughly worked out a couple of times during in the last one hundred years. See my *Logic and the Objectivity of Knowledge*, 143–66.

29 See Searle's remark to the effect that "philosophical quality varies inversely with the number of bibliographical references, and that no great work of philosophy ever contained a lot of footnotes," In *The Rediscovery of Mind*, xiv.

30 L. Susan Stebbing, *A Modern Introduction to Logic* (New York: Harper & Brothers, 1961), 214–15. Stebbing draws heavily on the three volume *Logic* of W. E. Johnson, especially Part II, Chapter I, §3.

31 Plantinga, *Warrant: the Current Debate*, 73.

32 For an account that covers much the same ground as this paper, but from within the framework and terminology of Husserl's strongly realist account of knowledge, see my paper, "Knowledge," in Barry Smith and David Woodruff Smith (eds.) *The Cambridge Companion to Husserl* (Cambridge: Cambridge University Press, 1995), 138–67.

3 The incompatibility of naturalism and scientific realism

Robert C. Koons

1. Introduction

Whenever philosophers bother to offer a defense for philosophical naturalism, they typically appeal to the authority of natural science. Science is supposed to provide us with a picture of the world so much more reliable and well-supported than that provided by any non-scientific source of information that we are entitled, perhaps even obliged, to withhold belief in anything that is not an intrinsic part of our our best scientific picture of the world. This scientism is taken to support philosophical naturalism, since, at present, our best scientific picture of the world is an essentially materialistic one, with no reference to causal agencies other than those that can be located within space and time.

This defense of naturalism presupposes a version of scientific realism: unless science provides us with objective truth about reality, it has no authority to dictate to us the form which our philosophical ontology and metaphysics must take. Science construed as a mere instrument for manipulating experience, or merely as an autonomous construction of our society, without reference to our reality, tells us nothing about what kinds of things really exist and act.

In this chapter, I will argue, somewhat paradoxically, that scientific realism can provide no support to philosophical naturalism. In fact, the situation is precisely the reverse: naturalism and scientific realism are incompatible.

Specifically, I will argue that (in the presence of certain well-established facts about scientific practice) the following three theses are mutually inconsistent:

1. Scientific realism
2. Ontological naturalism (the world of space and time is causally closed)
3. There exists a correct naturalistic account of knowledge and intentionality (representational naturalism)

By scientific realism, I intend a thesis that includes both a semantic and an epistemological component. Roughly speaking, scientific realism is the conjunction

of the following two claims:

1. Our scientific theories and models are theories and models of the real world, including its laws, as it is objectively, independent of our preferences and practices.
2. Scientific methods tend, in the long run, to increase our stock of real knowledge.

Ontological naturalism is the thesis that nothing can have any influence on events and conditions in space and time except other events and conditions in space and time. According to the ontological naturalist, there are no causal influences from things "outside" space: either there are no such things, or they have nothing to do with us and our world.

Representational naturalism is the proposition that human knowledge and intentionality are parts of nature, to be explained entirely in terms of scientifically understandable causal connections between brain states and the world. *Intentionality* is that feature of our thoughts and words that makes them *about* things, that gives them the capability of being true or false of the world.

I take philosophical naturalism to be the conjunction of the ontological and representational naturalism. The two theses are logically independent: it is possible to be an ontological naturalist without being a representational naturalist, and vice versa. For example, eliminativists like the Churchlands, Stich and (possible) Dennett are ontological naturalists who avoid being representational naturalists by failing to accept the reality of knowledge and intentionality. Conversely, a Platonist might accept that knowledge and intentionality are to be understood entirely in terms of causal relations, including, perhaps, causal connections to the forms, without being an ontological naturalist. I will argue that it is only the conjunction of the two naturalistic theses that is incompatible with scientific realism.

Many philosophers believe that scientific realism gives us good reason to believe both ontological naturalism and representational naturalism. I will argue, paradoxically, that scientific realism entails that either ontological naturalism or representational (or both) are false. I will argue that nature is comprehensible scientifically *only if* nature is *not* a causally closed system – only if nature is shaped by supernatural forces (forces beyond the scope of physical space and time).

My argument requires two critical assumptions:

PS: A preference for simplicity (elegance, symmetries, invariances) is a pervasive feature of scientific practice.

ER: Reliability is an essential component of knowledge and intentionality, on any naturalistic account of these.

2. The pervasiveness of simplicity

Philosophers and historians of science have long recognized that quasi-aesthetic

considerations, such as simplicity, symmetry, and elegance, have played a pervasive and indispensable role in theory choice. For instance, Copernicus's heliocentric model replaced the Ptolemaic system long before it had achieved a better fit with the data because of its far greater simplicity. Similarly, Newton's and Einstein's theories of gravitation won early acceptance due to their extraordinary degree of symmetry and elegance.

In his recent book, *Dreams of a Final Theory*, physicist Steven Weinberg included a chapter entitled "Beautiful Theories," in which he detailed the indispensable role of simplicity in the recent history of physics. According to Weinberg, physicists use aesthetic qualities both as a way of suggesting theories and, even more importantly, as a *sine qua non* of viable theories. Weinberg argues that this developing sense of the aesthetics of nature has proved to be a reliable indicator of theoretical truth.

> The physicist's sense of beauty is . . . supposed to serve a purpose – it is supposed to help the physicist select ideas that help us explain nature.[1]. . . we demand a simplicity and rigidity in our principles before we are willing to to take them seriously.[2]

For example, Weinberg points out that general relativity is attractive, not just for its symmetry, but for the fact that the symmetry between different frames of reference requires the existence of gravitation. The symmetry built into Einstein's theory is so powerful and exacting that concrete physical consequences, such as the inverse square law of gravity, follow inexorably. Similarly, Weinberg explains that the electroweak theory is grounded in an internal symmetry between the roles of electrons and neutrinos.

The simplicity that physicists discover in nature plays a critical heuristic role in the discovery of new laws. As Weinberg explains,

> Weirdly, although the beauty of physical theories is embodied in rigid, mathematical structures based on simple underlying principles, the structures that have this sort of beauty tend to survive even when the underlying principles are found to be wrong. . . . We are led to beautiful structures by physical principles, but the beauty sometimes survives when the principles themselves do not.[3]

For instance, Dirac's 1928 theory of the electron involved an elegant formalism. Dirac's theory led to the discovery of the positron, and the mathematics of Dirac's theory has survived as an essential part of quantum field theory, despite the fact that Dirac's approach to reconciling quantum mechanics and relativity was wrong.[4] Similarly, mathematicians' pursuit of elegant mathematical theories has regularly anticipated the needs of theoretical physicists. The theory of curved space was developed by Gauss and Riemann before it was needed by Einstein, and group theory

antedated its use in the theory of internal symmetry principles in particle physics.[5]

Weinberg notes that the simplicity that plays this central role in theoretical physics is "not the mechanical sort that can be measured by counting equations or symbols."[6] The recognition of this form of beauty requires an act of quasi-aesthetic judgment. As Weinberg observes,

> There is no logical formula that establishes a sharp dividing line between a beautiful explanatory theory and a mere list of data, but we know the difference when we see it.[7]

In claiming that an aesthetic form of simplicity plays a pervasive and indispensable role in scientific theory choice, I am not claiming that the aesthetic sense involved is innate or apriori. I am inclined to agree with Weinberg in thinking that "the universe acts as a random, inefficient and in the long-run effective teaching machine. . . ."[8] We have become attuned to the aesthetic deep structure of the universe by a long process of trial and error, a kind of natural selection of aesthetic judgments. As Weinberg puts it,

> Through countless false starts, we have gotten it beaten into us that nature is a certain way, and we have grown to look at that way that nature is as beautiful. . . . Evidently we have been changed by the universe acting as a teaching machine and imposing on us a sense of beauty with which our species was not born. Even mathematicians live in the real universe, and respond to its lessons.[9]

Nonetheless, even though we have no reason to think that the origin of our aesthetic attunement to the structure of the universe is mysteriously prior to experience, there remains the fact that experience has attuned us to *something*, and this something runs throughout the most fundamental laws of nature. Behind the blurrin' and buzzin' confusion of data, we have discovered a *consistent* aesthetic behind the various fundamental laws. As Weinberg concludes,

> It is when we study truly fundamental problems that we expect to find beautiful answers. We believe that, if we ask why the world is the way it is and then ask why that answer is the way it is, at the end of this chain of explanations we shall find a few simple principles of compelling beauty. We think this in part because our historical experience teaches us that as we look beneath the surface of things, we find more and more beauty.
>
> Plato and the neo-Platonists taught that the beauty we see in nature is a reflection of the beauty of the ultimate, the nous. For us, too, the beauty of present theories is an anticipation, a premonition, of the beauty of the final theory. And, in any case, we would not accept any theory as final unless it were

beautiful.[10] This capacity for "premonition" of the final theory is possible only because the fundamental principles of physics share a common bias toward a specific, learnable form of simplicity.

3. The centrality of reliability to representational naturalism

The representational naturalist holds that knowledge and intentionality are entirely natural phenomena, explicable in terms of causal relations between brain states and the represented conditions. In the case of knowledge, representational naturalism must make use of some form of reliability. The distinction between true belief and knowledge turns on epistemic norms of some kind. Unlike Platonists, representational naturalists cannot locate the basis of such norms in any transcendent realm. Consequently, the sort of *rightness* that qualifies a belief as knowledge must consist in some relation between the actual processes by which the belief is formed and the state of the represented conditions. Since knowledge is a form of success, this relation must involve a form of reliability, an objective tendency for beliefs formed in similar ways to represent the world accurately.

A representational naturalist might make use, as do Dretske, Papineau and Millikan, of teleological properties, so long as these are taken to consist in a set of causal and historical relations. Knowledge could then be identified with true beliefs formed by processes whose proper functions are fulfilled in normal circumstances. However, this teleological account also connects knowledge with reliability, since the proper function of belief-forming processes is to form true beliefs, so the sort of process which fulfills this proper function must be a reliable one.

Thus, if representational naturalism is combined with epistemic realism about scientific theories, the conjunction of the two theses entails that our processes of scientific research and theory choice must reliably converge upon the truth.

A naturalistic account of intentionality must also employ some notion of reliability. The association between belief-states and their truth-conditions must, for the representational naturalist, be a matter of some sort of natural, causal relation between the two. This association must consist in some sort of regular correlation between the belief-state and its truth-condition under certain conditions (the "normal" circumstances for the belief-state).

For example, according to Papineau, beliefs have teleological purposes, and these purposes fix their truth conditions, since "beliefs are true when they fulfill their purpose of co-varying with the relevant circumstances."[11] This co-variation of representation and represented condition is what gives the capacity for belief is biological value. "According to the natural-selection story it is the fact that a belief-type 'typically' obtains in certain circumstances that will explain our having it in our repertoire. . .".[12] This regular association of belief-type and truth-conditions, and the biological purposes which the association serves, provide exactly the kind of naturalistic explication of intentionality that the representational naturalist requires.

This regular association is a form of reliability. As Fodor observed:

> . . . we shall still have this connection between the etiology of representations and their truth values: representations generated in teleologically normal circumstances must be *true*.[13]

This reliability is only a conditional reliability: reliability under teleologically *normal* circumstances. This condition provides the basis for a distinction between knowledge and true belief: an act of knowledge that *p* is formed by processes that reliably track the fact that *p* in the actual circumstances, whereas a belief that *p* is formed by processes that would reliably track *p* in normal circumstances.

It is possible for our reliability to be lost. Conditions can change in such a way that teleologically normal circumstances are no longer possible. In such cases, our beliefs about certain subjects may become totally unreliable.

> It is the *past* predominance of true belief over false that is required.... [This] leaves it open that the statistical norm from now on might be falsity rather than truth. One obvious way in which this might come about is through a change in the environment.[14]

In addition, there may be specifiable conditions that occur with some regularity in which our belief-forming processes are unreliable.

> . . . this link is easily disrupted. Most obviously, there is the point that our natural inclinations to form beliefs will have been fostered by a limited range of environments, with the result that, if we move to new environments, those inclinations may tend systematically to give us false beliefs. To take a simple example, humans are notoriously inefficient of judging sizes underwater.[15]

Finally, the reliability involved may not involve a high degree of probability. The correlation of belief-type and represented condition does not have to be close to 1. As Millikan has observed, "it is conceivable that the devices that fix human beliefs fix true ones not on average, but just often enough."[16] For example, skittish animals may form the belief that a predator is near on the basis of very slight evidence. This belief will be true only rarely, but it must have a better-than-chance probability of truth under normal circumstances, if it is to have a representational function at all.

Thus, despite these qualifications, it remains the case that a circumscribed form of reliable association is essential to the naturalistic account of intentionality. The reliability is conditional, holding only under normal circumstances, and it may be minimal, involving a barely greater-than-chance correlation. Nonetheless, the

representational naturalist is committed to the existence of a real, objective association of the belief-state with its corresponding condition.

4. Proof of the incompatibility

I claim that the triad of scientific realism (SR), representational naturalism (RN), and ontological naturalism (ON) is inconsistent, given the theses of the pervasiveness of the simplicity criterion in our scientific practices (PS) and the essentiality of reliability as a component of naturalistic accounts of knowledge and intentionality. The argument for the inconsistency proceeds as follows.

1. SR, RN and ER entail that scientific methods are reliable sources of truth about the world.

As I have argued, a representational naturalist must attribute some form of reliability to our knowledge- and belief-forming practices. A scientific realist holds that scientific theories have objective truth-conditions, and that our scientific practices generate knowledge. Hence, the combination of scientific realism and representational naturalism entails the reliability of our scientific practices.

2. From PS, it follows that simplicity is a reliable indicator of the truth about natural laws.

Since the criterion of simplicity as a *sine qua non* of viable theories is a pervasive feature of our scientific practices, thesis 1 entails that simplicity is a reliable indicator of the truth (at the very least, a better-than-chance indicator of the truth in normal circumstances).

3. Mere correlation between simplicity and the laws of nature is not good enough: reliability requires that there be some causal mechanism connecting simplicity and the actual laws of nature.

Reliability means that the association between simplicity and truth cannot be coincidental. A regular, objection association must be grounded in some form of causal connection. Something must be causally responsible for the bias toward simplicity exhibited by the theoretically illuminated structure of nature.

4. Since the laws of nature pervade space and time, any such causal mechanism must exist outside space time.

By definition, the laws and fundamental structure of nature pervade nature. Anything that causes these laws to be simple, anything that imposes a consistent aesthetic upon them, must be supernatural.

5 Consequently, ON is false.

The existence of a supernatural cause of the simplicity of the laws of nature is obviously inconsistent with ontological naturalism. Hence, one cannot consistently embrace naturalism and scientific realism.

Frank Ramsey[17] and David Lewis[18] have proposed an account of the nature of natural law that would dispose of any need to explain the reliability of simplicity as an indicator of genuine lawhood. Their account simply identifies the laws of nature with the axioms of the best theory of the world, where *best* is cashed out in terms of such virtues as simplicity, strength and fit with the empirical data. Hence, it becomes an analytic truth that simplicity is a criterion of the lawfulness of a confirmed generalization.

In *Realism Regained*[19] I argue that the Ramsey/Lewis account of laws is incompatible with scientific realism (as I defined it) and fails adequately to explain the changing, dynamic nature of our standards of simplicity as we learn by experience what sort of aesthetic properties the laws of nature really have. On the Ramsey/Lewis account, facts about the laws of nature are not independent of our practices and preferences, and so lack the kind of objectivity required by scientific realism. In addition, Ramsey and Lewis cannot give a causal theory of scientific knowledge or intentionality, since, on their view, laws of nature are not the kinds of things that can enter into cause-and-effect relations with our cognitive faculties.

5. Papineau and Millikan on scientific realism

David Papineau and Ruth Garrett Millikan are two thoroughgoing naturalists who have explicitly embraced scientific realism. If the preceding argument is correct, this inconsistency should show itself somehow in their analyses of science. This expectation is indeed fulfilled. For example, Papineau recognizes the importance of simplicity in guiding the choice of fundamental scientific theories. He also recognizes that his account of intentionality entails that a scientific realist must affirm the reliability of simplicity as a sign of the truth. Nonetheless, he fails to see the incompatibility of this conclusion with his ontological naturalism. Here is the relevant passage:

> . . . it is plausible that at this level the inductive strategy used by physicists is to ignore any theories that lack a certain kind of *physical simplicity*. If this is right, then this inductive strategy, when applied to the question of the general constitution

> of the universe, will inevitably lead to the conclusion that the universe is composed of constituents which display the relevant kind of physical simplicity. And then, once we have reached this conclusion, we can use it to explain why this inductive strategy is reliable. For if the constituents of the world are indeed characterized by the relevant kind of physical simplicity, then a methodology which uses observations to decide between alternatives with this kind of simplicity will *for that reason* be a reliable route to the truth.[20]

In other words, so long as we are convinced that the laws of nature *just happen to be* simple in the appropriate way, we are entitled to conclude that our simplicity-preferring methods were *reliable* guides to the truth. However, it seems clear that such a retrospective analysis would instead reveal that we succeeded by sheer, dumb luck.

By way of analogy, suppose that I falsely believed that a certain coin was two-headed. I therefore guess that all of the first six flips of the coin will turn out to be heads. In fact, the coin is a fair one, and, by coincidence, the five of the first six flips did land heads. Would we say in this case that my assumption was a reliable guide to the truth about these coin flips? Should we say that its reliability was 5/6? To the contrary, we should say that my assumption led to very unreliable predictions, and the degree of success that I achieved was due to good luck, and nothing more.

Analogously, if it is a mere coincidence that the laws of nature share a certain form of aesthetic beauty, then our reliance upon aesthetic criteria in theory choice is not in any sense reliable, not even minimally reliable, not even reliable in ideal circumstances. When we use the fact that we have discovered a form of "physical simplicity" in law *A* as a reason for preferring theories of law *B* which have the same kind of simplicity, then our method is reliable only if there is some causal explanation of the repetition of this form of simplicity in nature. And this repetition necessitates a supernatural cause.

Papineau recognizes that we do rely on such an assumption of the repetition of simplicity:

> The account depends on the existence of certain general features which characterize the true answers to questions of fundamental physical theory. Far from being knowable *a priori*, these features may well be counterintuitive to the scientifically untrained.[21]

Through scientific experience, we are "trained" to recognize the simplicity shared by the fundamental laws, and we use this knowledge to anticipate the form of unknown laws. This projection of experience from one law to the next is reliable only if there is some common cause of the observed simplicity.

Similarly, Millikan believes that nature has trained into us (by trial and error learning) certain "principles of generalization and discrimination"[22] that provide us with a solution to the problem of theoretical knowledge that was "elegant, supremely

general, and powerful, indeed, I believe it was a solution that cut to the very bone of the ontological structure of the world."[23] However, Millikan seems unaware of just how deep this incision must go. A powerful and supremely general solution to the problem of theory choice must reach a ground of the common form of the laws of nature, and this ground must lie outside the bounds of nature.

Papineau and Millikan might try to salvage the reliability of a simplicity bias on the grounds that the laws of nature are, although uncaused, brute facts, *necessarily* what they are. If they share, coincidentally, a form of simplicity and do so non-contingently, then a scientific method biased toward the appropriate form of simplicity will be, under the circumstances, a reliable guide to the truth.

There are two compelling responses to this line of defense. First, there is no reason to suppose that the laws of nature are necessary. Cosmologists often explore the consequences of models of the universe in which the counterfactual laws hold.

Second, an unexplained coincidence, even if that coincidence is a brute-fact necessity, cannot ground the reliability of a method of inquiry. A method is reliable only when there is a causal mechanism that explains its reliability. By way of illustration, suppose that we grant the necessity of the past: given the present moment, all the actual events of the past are necessary. Next, suppose that a particular astrological method generates by chance the exact birthdate of the first President of the United States. Since that date is now necessary, there is no possibility of the astrological method's failing to give the correct answer. However, if there is no causal mechanism explaining the connection between the method's working and the particular facts involved in Washington's birth, then it would be Pickwickian to count the astrological method as *reliable* in investigating this particular event.

Analogously, if the various laws of nature just happen, as a matter of brute, inexplicable fact, to share a form of simplicity, then, even if this sharing is a matter of necessity, using simplicity as a guide in theory choice should not count as reliable.

6. The Forster–Sober account of simplicity

In a recent paper,[24] Malcolm Forster and Elliott Sober offer a justification of the scientific preference for simplicity that seems to be compatible with scientific realism and yet which does not acknowledge any sense in which simplicity is a reliable indicator of the truth. If the Forster-Sober account provides an adequate explanation of the role of simplicity without any such reliable connection between simplicity and truth, then it would provide a serious challenge to the argument of the previous section. As Forster and Sober put it,

> In the past, the curve fitting problem has posed a dilemma: Either accept a realist interpretation of science at the price of viewing simplicity as an irreducible and *a prioristic* sign of truth and thereby eschew empiricism, or embrace some form of anti-realism. Akaike's solution to the curve fitting problem dismantles the

dilemma. It is now possible to be a realist and an empiricist at the same time.[25]

The issue for Forster and Sober is realism vs. empiricism, whereas for us it is realism vs. naturalism, but it would seem that analogous claims could be made on behalf of Akaike's solution. This solution is supposed to give the realist some reason for preferring simpler hypotheses that is independent of any supposed correlation between simplicity and truth.

The Akaike solution goes something like this. First, we must assume that all of our observations involve a certain amount of noise – that random observational error regularly occurs, and the the error values are normally distributed. We divide the possible hypotheses into a finite sequence of families, based on the degree of simplicity (measured by the number of parameters that are allowed to vary within the family). Instead of selecting the hypothesis that best fits the actual data, we instead look for a family of hypotheses with the best combination of goodness-of-fit and simplicity, and choose the best fitting hypothesis within that set.

The rationale for the Akaike criterion is the avoidance of *overfitting*. Since the actual data includes some unknown observational error, the curve that best fits the data is unlikely to be the true one. It will tend to fit the actual data better than the true curve, which is called the "overfitting" of the hypothesis to the data. Balancing goodness-of-fit with simplicity is supposed to mitigate this overfitting error. Consequently, the realist is given some reason to employ simplicity as a desideratum of theory choice without assuming any correlation between simplicity and truth.

Simpler, low-dimensional families are much smaller than the more complex, high-dimensional families. There are therefore two reasons why the more complex families are more likely to contain the hypothesis that best fits the data:

(a) Larger families generally contain curves closer to the truth than smaller families.
(b) *Overfitting:* the higher the number of adjustable parameters, the more prone the family is to fit to noise in the data.[26]

According to Forster and Sober, we want to favor a family of hypotheses if it contains a good fit to the data because of reason (a), but not if it contains one because of reason (b). What is needed is an estimate of the expected degree of overfitting associated with each family, given the actual data. Akaike demonstrated that, under certain special conditions, we can find an *unbiased estimator* of this special form of error. By subtracting the number of parameters that are allowed to vary within a family from a measure of the degree-of-fit of the best-fitting curve within that family (this measure is one of log-likelihood or, in special cases, the sum of squares), we can arrive at a *corrected* estimate of the degree of fit of the family to the truth, which Forster and Sober call the "expected predictive accuracy" of the family.[27] The Akaike criterion tells us to choose the best-fitting hypothesis within the family with the greatest expected predictive accuracy. In this way, we have both a definite rule for trading-off goodness-of-fit for

simplicity, and a plausible rationale for making the tradeoff.

There are several points to be made in response to this solution. First, it is not at all clear that the role of simplicity in the kind of curve-fitting practices Forster and Sober discuss is at all analogous to the role simplicity plays in our choice of fundamental physical theories. As Weinberg observed, the kind of *simplicity* that guides our choice of fundamental theories is not easily defined. It does not correspond directly to what Forster and Sober mean by the *simplicity* of a family of hypotheses, namely, the number of variable parameters in the corresponding equations.

Second, the technical results upon which Forster and Sober rely are quite limited in their scope of application, as I. A. Kieseppä has demonstrated.[28] The Aikake estimator of predictive accuracy is valid only when the space of hypotheses is carefully circumscribed. For example, it is valid when the space of hypotheses includes only polynomial equations, but invalid when it includes periodic functions, like the sine wave function.[29]

Third, the rationale for the Akaike criterion is incompatible with the reliabilist implications of combining scientific realism with representational naturalism. The sort of "scientific realism" that Forster and Sober have in mind is much less specific, implying only a concern with the truth of our scientific theories. Forster and Sober make no effort to demonstrate that reliance on the Akaike criterion leads reliably to the truth. Instead, they provide only a rationale that might reasonably motivate a realist to prefer simpler theories.

Finally, it is far from clear that even this rationale provides a basis for preferring simplicity that is genuinely independent of the reliability of simplicity as a sign of the truth. As has been pointed out by Kieseppä,[30] Scott De Vito,[31] and Andre Kukla,[32] the Akaike solution presupposes that a determinate conception of simplicity is a given. There is no objective, language- and representation-independent way of "counting the parameters" associated with a given curve. A linear curve is *naturally* thought of as having a single parameter, but this can easily be altered by redescribing the curve or altering the coordinate system. Sorting hypotheses into families by simplicity as we perceive it reflects a prior and unjustified preference for some hypotheses over others.

Forster and Sober might insist that the sorting of hypotheses into a hierarchy of families is entirely arbitrary or random. As they present the argument for the Akaike criterion, all that matters is that the hypotheses be sorted into a sequence of families in which the size of the families increases exponentially, and that this sorting *not* be done in an *ad hoc* fashion, in response to the actual data observed. Then, when we observe a relatively small family F with a hypothesis h showing a surprisingly good degree of fit to the data (surprising, that is, in light of the smallness of F), we are supposed to have good reason to believe that F has a high degree of predictive accuracy, and, therefore, that we have reason to prefer h over other hypotheses with better fit that happen to belong to much larger families. However, if it was entirely a matter of chance or caprice that h ended up in a small family, and its better-fitting competitors ended up in larger families, it is hard to see how h's good fortune provides us with any rational ground for preferring it.

To the contrary, the plausibility of the Akaike solution depends on our prior conviction that simpler hypotheses (as measured by mathematical conventions that have proved reliable at this very task) are disproportionately probable. What Forster and Sober give us is a principled way of weighing the two competing desiderata of simplicity and goodness of fit, but they do not provide us with a rationale for treating simplicity as a desideratum in the first place.

Consequently, Forster and Sober do not provide us with a way of escaping the conclusion that a reliabilist conception of scientific realism entails the reliability of simplicity as an indicator of the truth.

7. Pragmatic accounts of the simplicity criterion

A popular strategy for explaining the role of simplicity in scientific theorizing has been to appeal to a variety of pragmatic considerations. For example, Reichenbach argued that we favor simpler hypotheses because they are easier to represent, to make deductions from, and to use in calculations.[33] More recently, Peter Turney has argued that simpler hypotheses are more likely (given the presence of random observational error) to be repeatedly confirmed.[34]

However, these pragmatic justifications again sidestep the central issue, that of *reliability*. If our reliance on simplicity is unreliable, resulting in a bias toward simplicity that is not reflected in the constitution of nature, then we cannot combine scientific realism with representational naturalism.

A pragmatic justification of our scientific practice, when combined with representational naturalism, yields the conclusion that scientific theories must be interpreted non-representationally, either as mere instruments for generating empirical predictions, or as conventional constructs valid only for a local culture. Pragmatism, by eschewing any commitment to the objective reliability of scientific methods, cannot be combined with a naturalistic version of scientific realism.

8. Conclusion

Philosophical naturalism, then, can draw no legitimate support from the deliverances of natural science, realistically construed, since scientific realism entails the falsity of naturalism. If scientific theories are construed non-realistically, it seems that the status of ontology cannot be affected by the successes of natural science, nor by the form that successful theories in the natural sciences happen to take. If scientific anti-realism is correct, then the "manifest image" of the scientific world-view must not be taken as authoritative. Instead, that image is merely a useful fiction, and metaphysics is left exactly as it was before the advent of science.

Of course, naturalism as a metaphysical program existed before the development of modern science (Democritus, Epicurus, Lucretius) and presumably it would

survive the downfall of scientific realism. However, modern naturalists owe the rest of us a rational basis for their preferences that is independent of science.

In fact, the situation for the naturalist is even worse than I have described it. To the extent that the success of natural science provides support for scientific realism (in both its semantic and epistemic versions), to that extent it provides grounds for rejecting philosophical naturalism. Thus, conventional wisdom has the relationship between natural science and naturalism exactly backwards. In fact, the more successes natural science accumulates, the less plausible philosophical naturalism becomes.

There is a third thesis that is often included (especially since Quine) in the definition of naturalism: the continuity between the methods of philosophy and those of natural science, which we might call "metaphilosophical naturalism." Scientific antirealism, when combined with meta-philosophical naturalism, leads to the conclusion of philosophical antirealism, since philosophical theories are, according to metaphilosophical naturalism, merely a species of scientific theories.

This means that full-orbed naturalism (ontological + representational + metaphilosophical) is a self-defeating position. Full-orbed naturalism is a philosophical theory, and yet it entails philosophical anti-realism, which means that such theories cannot be known, and do not even purport to represent the world. Full-orbed naturalism cannot be true, since if it were true, it would entail that no philosophical theory (itself included) could be true.

Notes

1 Steven Weinberg, *Dreams of a Final Theory: The Scientist's Search for the Ultimate Laws of Nature* (New York: Vintage Books, 1993), 133.
2 Ibid., 148–9.
3 Ibid., 151–2.
4 Ibid., 151.
5 Ibid., 152.
6 Ibid., 134.
7 Ibid., 148–9.
8 Ibid., 158.
9 Ibid., 158–9.
10 Ibid., 165.
11 David Papineau, *Philosophical Naturalism* (Oxford: Blackwell, 1993), 177.
12 David Papineau, "Representation and Explanation," *Philosophy of Science* 51 (1984): 558.
13 Jerry Fodor, "Semantics, Wisconsin Style," *Synthese* 59 (1984): 247.
14 Papineau, "Representation and Explanation," 558.
15 Papineau, *Philosophical Naturalism,* 100.
16 Ruth Garrett Millikan, "Biosemantics," *Journal of Philosophy* 86 (1989): 289.
17 F.P. Ramsey, *Philosophical Papers* (Cambridge: Cambridge University Press, 1990).
18 David Lewis, "Humean Supervenience Debugged," Mind 103 (1994): 473–90.
19 Robert C. Koons, *Realism Regained: An Exact Theory of Causation, Teleology and the Mind* (New York: Oxford University Press, 2000).
20 Papineau, *Philosophical Naturalism*, 166.

21 Ibid., 166.
22 Millikan, "Biosemantics," 292.
23 Ibid., 294.
24 Malcolm Forster and Elliot Sober, "How to Tell When Simpler, More Unified, or Less *Ad Hoc* Theories Will Provide More Accurate Predictions," *British Journal for the Philosophy of Science* 46 (1995): 248–52.
25 Ibid., 28.
26 Ibid., 8.
27 Ibid., 2.
28 I.A. Kieseppä, "Akaike Information Criterion, Curve-fitting and the Philosophical Problem of Simplicity", *British Journal for the Philosophy of Science* 48 (1997).
29 Ibid., 34–7.
30 Kieseppä, "Akaike Information Criterion."
31 Scott De Vito, "A Gruesome Problem for the Curve-Fitting Solution," *British Journal for the Philosophy of Science* 48 (1997): 391–6.
32 Andre Kukla, "Forster and Sober and the Curve-Fitting Problem," *British Journal for the Philosophy of Science* 46 (1995): 248–52.
33 Hans Reichenbach, "The Pragmatic Justification of Induction," in H. Feigl and W. Sellars (eds.), *Readings in Philosophical Analysis* (New York: Appleton-Century-Crofts), 305–27.
34 Peter Turney, "The Curve-Fitting Problem, a Solution," *British Journal for the Philosophy of Science* 41 (1990): 509–30.

Part 2

Ontology

4 Naturalism and the ontological status of properties

J. P. Moreland

Throughout the history of philosophy, many have advanced the idea that there is a connection between naturalism and nominalism. Specifically, many philosophers have argued that naturalism requires a nominalist rejection of the existence of properties construed along traditional realist lines as abstract, multiply-exemplifiable entities that non-spatially inhere in their instances.[1] Thus, in *Sophist* 246A-C, we read these words from Plato:

Stranger: What we shall see is something like a battle of gods and giants going on between them over their quarrel about reality.

Theaetetus: How so?

Stranger: One party is trying to drag everything down to earth out of heaven and the unseen, literally grasping rocks and trees in their hands, for they lay hold upon every stock and stone and strenuously affirm that real existence belongs only to that which can be handled and offers resistance to the touch. They define reality as the same thing as body, and as soon as one of the opposite party asserts that anything without a body is real, they are utterly contemptuous and will not listen to another word.

Theaetetus: The people you describe are certainly a formidable crew. I have met quite a number of them before now.

Stranger: Yes, and accordingly their adversaries are very wary in defending their position somewhere in the heights of the unseen, maintaining with all their force that true reality consists in certain intelligible and bodiless forms. In the clash of argument they shatter and pulverize those bodies which their opponents wield, and what those others allege to be true reality they call, not real being, but a sort of moving process of becoming. On this issue an interminable battle is always going on between the two camps.

In the contemporary setting, Howard Robinson remarks that "materialist theories are incompatible with realist theories of universals. The tie between nominalism and materialism is an ancient one."[2] Since most philosophers currently see physicalism (which for our purposes can be used synonymously with materialism) as the best form of naturalism, Robinson's remark applies with equal force to a widely accepted

contemporary version of naturalism to be described below. Along similar lines, Reinhardt Grossmann has argued that naturalists are at war with what he calls ontologists.[3] According to Grossmann, the universe is the spatio-temporal totality of physical entities and the world includes every existent whatever, including non-spatio-temporal abstract entities such as properties. Naturalists deny the world and only believe in the universe; ontologists like Grossmann accept the world. Wilfrid Sellars claimed that "a naturalist ontology must be a nominalist ontology."[4] Elsewhere, Sellars argued that a nominalist analysis of predication is the ". . . very foundation of a naturalist ontology."[5]

However, not all philosophers have agreed with this view of the relationship between naturalism and nominalism. For example, Colin McGinn claims that as long as the naturalist gives an account of the instantiation of all properties in terms of the instantiation of physical properties taken as straightforward physical facts, (s)he is free to take properties themselves as non-physical, abstract entities.[6]

Currently, debates about naturalism and abstract objects usually focus on mathematical entities, e.g. Quine is called a Platonist for embracing sets, Hartry Field adopts a fictional instrumentalist account of mathematical truth as a requirement for expressing a consistent physicalist naturalism, and Penelope Maddy tries to work out a version of realism in mathematics that is consistent with naturalism as she conceives it.[7] As interesting as the debate about mathematical entities is, for three reasons, I think it is a mistake to focus the dialog about naturalism and abstract objects on mathematical entities instead of on properties. For one thing, without begging any questions against nominalists, the existence and nature of properties is just more pre-analytically evident than the existence and nature of numbers. Thus, pre-analytic intuitions allow for more flexibility for analyses of numbers than for properties. For example, if a philosopher identifies the number two with a set because the identity conditions of the latter are unproblematic, there is more prima facie plausibility in that move than for a similar strategy for treating the property of being red.

Second, properties are every bit as central to scientific theory as mathematical entities, so no advantage is gained for those who wish to give pride of place to scientific discourse in ontology by focusing on mathematics.

Third, and most importantly, properties are far more critical for fleshing out an adequate ontological framework than mathematical entities. Virtually any ontological issue will require properties or at least property-talk for its expression, but the same cannot be said for mathematics. Indeed, some solutions to the ontological status of mathematical entities employ properties (or at least property-talk). For example, Maddy attempts to work out a version of realism consistent with her conception of the constraints that follow from naturalism (e.g. rejection of first philosophy, justification of entities because they are required by our best scientific theories, acceptance of

entities only if knowledge of them can be harmonized with a naturalist view of humans as physical cognizing entities).[8] For Maddy, numbers are properties of sets and sets are located at the place where the physical objects which are members of the sets are. Now Maddy's solution requires at least property-talk and she explicitly leaves open the realist/nominalist debate about properties. For our purposes, I note simply that the debate about properties is more fundamental to Maddy's ontology than is her position on numbers.

It would seem, then, that debates about the ontological status of properties are central to the task of analyzing the adequacy of different forms of naturalism. There are four strategies that naturalists can employ in dealing with properties: (1) Accept a traditional realist construal of properties and claim that this is fully consistent with a widely accepted version of contemporary naturalism. (2) Accept a traditional realist construal of properties and take this to require an abandonment of a widely accepted version of contemporary naturalism. (3) Reject a traditional realist and accept a modified construal of properties which is acceptable within the constraints of contemporary naturalism. (4) Reject the existence of properties in favor of extreme nominalism.

The thesis I wish to defend is (2). If properties exist and are what traditional realists take them to be, then a widely accepted contemporary version of naturalism is false. (4) is the path taken by naturalists such as Wilfrid Sellars. I shall not look at this strategy due to space considerations and because I think the arguments against extreme nominalism are persuasive and widely known.[9] I will list but for two reasons not develop a set of considerations against (1). First, in my opinion those naturalists who take this option often do so apparently without being aware of the fundamental issues involved in embracing this alternative. I believe this often happens because these issues are matters of so-called first philosophy and many naturalists seem uninterested in or simply reject topics in first philosophy. (3) brings these issues to center stage and by focusing on them, I hope to show how crucial they are to debates about the adequacy of naturalism. Second, my interests simply lie in the current debate about (3), so if I can show it to be problematic, I can at least raise problems for those naturalists who agree with me that (3) is the key area of debate.

In what follows, I shall first, compare three schools of thought regarding the ontological status of properties and clarify the traditional realist approach; second, characterize three crucial elements that constitute a widely accepted contemporary version of naturalism and show why traditional realist depictions of properties are not a plausible option for naturalism so conceived; third, analyze and criticize the two most powerful naturalist attempts to reconcile naturalism with an acceptance of properties – the views of Armstrong and Campbell.

Three schools of thought regarding the ontological status of properties

There are three major schools of thought regarding the nature of properties. They can be clarified by focusing on attribute-agreement. Suppose we have before us two red and round spots. Suppose further that each spot has the "same" shade of red and the same roundness. Let us call the two spots Socrates and Plato. Let us also use red_1 and red_2 to stand for the redness of Socrates and the redness of Plato, respectively.

Attribute-agreement can be interpreted in three general ways. First, there is extreme nominalism. There are several versions of extreme nominalism, but all exclude properties as they are construed by the realists or nominalists, The extreme nominalist is committed to a reductive analysis of attributes:

a has the attribute F, if and only if, Q.

Different versions of extreme nominalism will spell out Q in different ways. A predicate extreme nominalist may spell out Q as "the predicate 'F' is true of a." A class extreme nominalist may state Q as "a is a member of the class of F-things." A concept extreme nominalist would substitute for Q something like "a falls under the concept F." My purpose is not to delineate all the varieties of extreme nominalism. I merely note that the fundamental feature of this account of attribute-agreement and exemplification is its denial that attributes form an additional category of being distinct from the things that have them (unless, of course, a new category other than that of quality is introduced and to which qualities are reduced, e.g. predicates, classes, concepts, etc.).

The second major interpretation of attribute-agreement is called nominalism. A nominalist acknowledges the existence of properties but denies that attribute-agreement is to be explained along realist lines where properties are taken to be universals. The nominalist denies that the *redness* of Socrates is numerically identical to the *redness* of Plato. Socrates and Plato may both have a determinate shade of color that is "exactly alike." But the two do not share the same numerically identical quality. Each has a particular entity that is not multiply exemplifiable, a little red. Quality-instances construed along nominalist lines have various labels: Among them are "tropes,"[10] "abstract particulars,"[11] "perfect particulars,"[12] "cases,"[13] "aspects,"[14] "unit properties,"[15] "property-instances,"[16] and "moments."[17]

Finally, the third type of interpretation of attribute-agreement is called the realist view. There are different varieties of realism. For example, Aristotelian realists disagree with Platonic realists over the question of the existence of uninstantiated universals. However, traditional realists like Reinhardt Grossmann hold universals to be non-spatio-temporal abstract entities which are constituents "in" their instances.[18]

To elaborate on a traditional realist understanding of qualities and quality-instances, recall our round, red spot, Socrates. According to traditional realism, a universal is a multiply exemplifiable entity which is a numerically identical entity in each of its instances. Thus, instances are not simples for a traditional realist; they are complex entities. Now consider the following three sentences and the states of affairs they describe.

(1) Socrates (a red, round spot) is red.
(2) Red_1 is red.
(3) This (individuator, e.g. a bare particular) is red.

The state of affairs described by (1) is best understood as follows. Socrates has a property-instance in it, red_1, and red_1 is red. The relation between red_1 and Socrates is a sort of part/whole relation. Let us call this the "moment/whole" relation, to borrow Husserl's terminology.[19] Observe that (1) is grounded, at least in part, in (2). The "is" in (1) is grounded in two different "ways of being in" – the way a moment is in a whole (Socrates) and the "way of being in" expressed in (2).

The state of affairs described in (2) contains a second kind of part/whole relation. (2) states that the universal, redness, is a part of the whole, red1 in the sense that redness is one of the constituents that is *in* red_1. This is not a spatial sense of *in*. Let us call this "way of being in" the relation of "being an essential property-constituent of". Two things should be said about this relation as it is found in (2). First, this is a part/whole relation that differs from the "moment/ whole" relation of (1) in at least one important way. If b stands to a in such a way that b "is an essential property-constituent of" a, then the following two principles hold between a and b that fail to obtain in the moment/whole relation: (1) b transcends a in the sense that b is necessary for a's existence but not vice versa. Moreover, b enters into the being of a but not vice versa. Here b is redness and a is red1. Redness can exist without red_1, and not vice versa. Red_1 could be destroyed but redness would still exist unchanged. Thus, there is a modal distinction between a universal and its instances. When a universal is exemplified, the universal is modified and constitutes the essence of its instances which, in turn, are complex, dependent particulars. The same could not be said for the "moment/whole" relation in (1) between Socrates and red_1. (2) B is the essence of a; that is, b answers the question "What is a an instance of?" Redness is the essence of red_1. Red_1 is an instance of redness in that redness is immanent within red_1 as its informing essence.

Second, this last point suggests an implication of (2). (2) entails an "is" of classification. It implies that red_1 is to be placed in the class of red instances, where "class" refers to a group of entities that literally share some entity in common. An "is"

of classification is more clearly expressed in

(2′) Red_1 is a red.

This sentence expresses the fact that red1 is identical to some red instance in the class of red instances. Moreover, (2) and (2′) are grounded in a more basic relation which founds the inclusion of red_1 in the class of red instances. This more basic relation is expressed in (3).

The state of affairs described by (3) contains an "is" of exemplification or predication. I use these two as synonyms here. The "is" in (1) and (2) are expressions of predication in a loose and popular sense and not in a strict, philosophical sense. This is because the "is"'s in (1) and (2) are grounded in a more basic relation. (1) is grounded (in part) in (2), and (2) is grounded in (3). The "is" in (3) refers to a basic, undefined, irreducible nexus of predication. It expresses the fact that the universal, redness, is connected to a bare particular (or, perhaps, some other individuator) in red1 by the nexus of exemplification. This nexus of predication is inhomogeneous in the sense that it relates two entities that differ categorically, a universal and an individuator. So redness is in an individuator (which is in red_1) in that redness is predicated of or "tied to" that individuator.

This, then, is a sketch of a traditional realist assay of property-instances which (1) maintains true universals, (2) grounds the individuation of property-instances by making them complex entities with an individuating constituent in them, and (3) clarifies three important states of affairs and their associated "ways of being in." For traditional realists, neither the universal nor the exemplification nexus are spatio-temporal entities. Moreover, in the case of Socrates and Plato, the exemplification nexus connects an abstract entity with a spatio-temporal one.

Johanna Seibt notes approvingly that Sellars argued several decades ago that "since the basic relation of a Platonist theory of predication, i.e. exemplification, cannot be defined in naturalist terms, a nominalist theory of predication proves to be '. . . the very foundation of a naturalist ontology'."[20] It may now be evident to the reader why Sellars thought this was so. If not, it should become evident once we get clear on the nature of a widely accepted version of contemporary naturalism to which Sellars subscribed.

Naturalism and the rejection of a traditional realist view of properties

Contemporary philosophical naturalism

Just exactly what are the central features of contemporary scientific naturalism?[21] As

with any widely accepted ideology, there will be different nuances given to naturalism by different thinkers. Nevertheless, it is both possible and desirable to give an accurate characterization of a specific form of philosophical naturalism (hereafter, simply naturalism or scientific naturalism) that is currently enjoying widespread acceptance. Roughly, naturalism is the view that the spatio-temporal universe of entities postulated by our best (or ideal) theories in the physical sciences is all there is. Scientific naturalism includes (1) different aspects of a naturalist epistemic attitude (e.g. a rejection of so-called first philosophy along with an acceptance of either weak or strong scientism); (2) an etiological account of how all entities whatsoever have come to be, constituted by an event causal story (especially the atomic theory of matter and evolutionary biology) described in natural scientific terms; and (3) a general ontology in which the only entities allowed are ones that bear a relevant similarity to those thought to characterize a completed form of physics.

The ordering of these three ingredients is important. Frequently, the naturalist epistemic attitude serves as justification for the naturalist etiology which, in turn, helps to justify the naturalist's ontological commitment. Moreover, naturalism seems to require a coherence among what is postulated in these three different areas of the naturalistic turn. For example, there should be a coherence among third-person scientific ways of knowing, a physical, evolutionary account of how our sensory and cognitive processes came to be, and an ontological analysis of those processes themselves. Any entities that are taken to exist should bear a relevant similarity to entities that characterize our best (or ideal) physical theories, their coming-to-be should be intelligible in light of the naturalist causal story, and they should be knowable by scientific means.

1. The naturalist epistemic attitude

As with much of modern philosophy, naturalism first and foremost is an expression of an epistemic posture, specifically, a posture called scientism. In the early 1960s, Wilfrid Sellars expressed this posture when he said that "in the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that it is not."[22] Steven Wagner and Richard Warner claim that naturalism is "the view that only natural science deserves full and unqualified credence."[23] Contemporary naturalists embrace either weak or strong scientism. According to the former, non-scientific fields are not worthless nor do they offer no intellectual results, but they are vastly inferior to science in their epistemic standing and do not merit full credence. According to the latter, unqualified cognitive value resides in science and in nothing else. Either way, naturalists are extremely skeptical of any claims about reality that are not justified by scientific methodology.

Naturalists believe that they are justified in this posture because of the success of science vis a vis other fields of inquiry. Also, some naturalists justify this standpoint by

appealing to the unity of science, though this argument is employed less frequently today than it was a few decades ago. For example, Roy Bhaskar asserts that "*naturalism* may be defined as the thesis that there is (or can be) an essential unity of method between the natural and the social sciences."[24] And as John Searle notes, since for these naturalists science exhausts what we can know, then belief in the unity of science turns out to be a belief in the unity of all knowledge because it is scientific knowledge:

> Every fact in the universe is in principle knowable and understandable by human investigators. Because reality is physical, and because science concerns the investigation of physical reality, and because there are no limits on what we can know of physical reality, it follows that all facts are knowable and understandable by us.[25]

For the naturalist, the exhaustive nature of scientific knowledge entails that the only explanations that count are scientific explanations.

We have seen that scientism is the core epistemic posture of the contemporary naturalist. From this core commitment, at least two philosophical theses follow that in one way or another elaborate the epistemic and methodological constraints for philosophy that are part of taking the naturalistic turn. First, there is no such thing as first philosophy. According to David Papineau, there is a continuity between philosophy and natural science:

> . . . the task of the philosophers is to bring coherence and order to the total set of assumptions we use to explain the natural world. The question at issue is whether *all* philosophical theorizing is of this kind. Naturalists will say that it is. Those with a more traditional attitude to philosophy will disagree. These traditionalists will allow, of course, that some philosophical problems, problems in *applied* philosophy, as it were, will fit the above account. But they will insist that when we turn to 'first philosophy', to the investigation of such fundamental categories as thought and knowledge, then philosophy must proceed independently of science. Naturalists will respond that there is no reason to place first philosophy outside of science.[26]

A second, and closely related thesis follows from the fact that "the philosopher who wants to regard human beings and mental phenomena as part of the natural order [must] explain intentional relations in naturalistic terms."[27] The naturalist will hold to an externalist theory of knowledge or justification (or warrant). This is true for two main reasons, one positive and one negative. The positive reason is that "A scientific or naturalistic account of [human beings and their mental states] must be a causal

account."[28] Externalist theories in epistemology either implicitly (e.g. reliabilist theories) or explicitly (e.g. causal accounts) center on the notion of causality.

The negative reason is this: The central notion of "internal" for the internalist is an irreducibly mental one that makes reference to such things as being a self-presenting property, being a mode of conscious, being directly present to the cognizer's awareness, cognitive accessibility, being internal to one's point of view, first person introspection, etc. Arguably, all these notions are mental. Moreover, they are irreducible and cannot be analyzed merely as functional states if the perspective is to remain an internalist one because if these are reduced to (or replaced by) physicalist (causal/functional) notions, then what is distinctive to internalism will fall out and the view will collapse into externalism.[29] Why? Because neither the properties just mentioned nor the way that they are internal to epistemic subjects (e.g. by being self-presenting, being available to private, first person access) is identifiable with (1) a particular, property, or relation that belongs to the ontology of physics or (2) a functional state or a physical realizer of a functional state.

Regarding the way they are internal to the subject on internalist theories, it is interesting to note that the most frequently employed candidates for the characteristic mark of (at least most) mental states not only serve as ontological criteria, they also capture crucial elements for internalist epistemic theories. While it is difficult to give a characteristic mark of the mental (or the physical!) in terms that command universal assent, nevertheless, most agree that the most likely candidate is this: a property (or event) is mental just in case it is such that it is self-presenting or that to which the subject has private or first-person direct access or that which could not exist unless there were sentient subjects. However, it has been hard to sustain the claim that this characteristic mark could be true of a physical particular, property, or relation. Internalist properties and their relationship to epistemic subjects, as these actually function in internalist epistemological theories, are just not reducible to natural, physical, causal/functional entities. Thus, if the best version of naturalism is strict physicalism, then the best naturalist theory of knowledge or justification (or warrant) will be externalist.

We have looked at a number of philosophers who express different aspects of the epistemic attitude that constitutes naturalism. Let us now turn to an overview of the naturalist's view of how things came to be.

2. The naturalist grand story

The naturalist has an account of how all things whatever came to be. Let us call this account the Grand Story. The details of the Grand Story need not concern us here. Suffice it to say that some version of the Big Bang is the most reasonable view currently available. On this view, all of reality – space, time, and matter – came from the original "creation" event and the various galaxies, stars, and other heavenly bodies eventually

developed as the expanding universe went through various stages. On at least one of those heavenly bodies – earth – some sort of pre-biotic soup scenario explains how living things came into being from non-living chemicals. And the processes of evolution, understood in either neo-Darwinian or punctuated equilibrium terms, gave rise to all the life forms we see including human beings. Thus, all organisms and their parts exist and are what they are because they contributed to (or at least did not hinder) the struggle for reproductive advantage, more specifically, because they contributed to the tasks of feeding, fighting, fleeing, and reproducing.

There are four important things to note about the Grand Story. First, the core of the Grand Story is the atomic theory of matter and evolutionary theory. If we take John Searle to be representative of naturalists here, this means that causal explanations, specifically, bottom-up but not top-down causal explanations, are central to the (alleged) explanatory superiority of the Grand Story.[30] Second, it is an expression of a naturalist version of philosophical monism according to which everything that exists or happens in the world is susceptible to explanations by natural scientific methods. Whatever exists or happens in the world is natural in this sense. Prima facie, the most consistent way to understand naturalism in this regard is to see it as entailing some version of strict physicalism: everything that exists is fundamentally matter, most likely, elementary particles (whether taken as points of potentiality, centers of mass/energy, or units of spatially extended stuff/waves), organized in various ways according to the laws of nature. By keeping track of these particles and their physical traits we are keeping track of everything that exists. No non-physical entities exist, including emergent ones. This constitutes a strict sense of physicalism. When naturalists venture away from strict physicalism, however, they still argue that additions to a strict physicalist ontology must be depicted as rooted in, emergent from, dependent upon the physical states and events of the Grand Story.

Third, a naturalist depicts living organisms as property-things or ordered aggregates. This follows from the two key elements of the Grand Story: the atomic theory of matter (and the rejection of any form of vitalism) and evolutionary biology (where macro-features – phenotypes – are explained by causal mechanisms at the micro-level – genotypes – and in terms of so-called functional explanations in terms of the constrains of reproductive advantage). The conjunction of these two elements implies two things about living creatures. To begin with, they are not genuine substances understood, in the classic sense, as essentially characterized particulars (1) whose parts stand in internal relations to the substantial whole; (2) which sustain absolute sameness through change; and (3) which have a nature possessed by all members of their natural kinds. Darwin's theory of evolution has made belief in, e.g. human substances with human natures, though logically possible, nevertheless, quite implausible. As E. Mayr has said:

> The concepts of unchanging essences and of complete discontinuities between every eidos (type) and all others make genuine evolutionary thinking impossible. I agree with those who claim that the essentialist philosophies of Aristotle and Plato are incompatible with evolutionary thinking.[31]

This belief has, in turn, lead thinkers like David Hull to make the following observation:

> The implications of moving species from the metaphysical category that can appropriately be characterized in terms of 'natures' to a category for which such characterizations are inappropriate are extensive and fundamental. If species evolve in anything like the way that Darwin thought they did, then they cannot possibly have the sort of natures that traditional philosophers claimed they did. If species in general lack natures, then so does *Homo sapiens* as a biological species. If *Homo sapiens* lacks a nature, then no reference to biology can be made to support one's claims about 'human nature.' Perhaps all people are 'persons,' share the same 'personhood,' etc., but such claims must be explicated and defended *with no reference to biology*. Because so many moral, ethical, and political theories depend on some notion or other of human nature, Darwin's theory brought into question all these theories. The implications are not entailments. One can always dissociate *'Homo sapiens'* from 'human being,' but the result is a much less plausible position.[32]

Thus, living organisms are not genuine substances for the naturalist. Rather, they are property-things. I mean by a "property-thing" an ordered aggregate, a set of parts put into external relations by some ordering principle (e.g. a law of nature) to form a whole. For our purposes, it is important to observe that given the naturalist etiology in general and evolutionary theory in particular, naturalists like Hull and Mayr conclude that certain abstract objects, namely, natural kind essences, either do not exist or they are not exemplified.

Finally, while some naturalists eschew questions about the nature of existence itself, others have formulated a definition of existence based on a naturalist epistemology and consistent with the Grand Story. Thus, Bruce Aune defines *a exists* as "a belongs to the space-time-causal system that is our world. Our world is, again, that system of (roughly) causally related objects. . .".[33] Along similar lines, D. M. Armstrong says that for any entities, the following question settles the issue of whether or not those entities can be said to exist: "Are these entities, or are they not, capable of action upon the spatio-temporal system? Do these entities, or do they not, act in nature?"[34] Daniel Dennett claims that when we are trying to find out whether or not some entity

like the self exists, what we must do is locate the entity within the causal fabric.[35] Keith Campbell applies the same reasoning to the question of the existence of emergent entities like social characteristics by claiming that the test of their existence turns on their ability to exhibit independent causality because "power has been recognized as the mark of being."[36] Finally, Jaegwon Kim says "Causal powers and reality go hand in hand. To render mental events causally impotent is as good as banishing them from our ontology."[37]

3. The naturalist ontology

In order to characterize a naturalist ontology, we may distinguish global vs. local naturalism and weak vs. strong naturalism. Roughly, global naturalism is the view that the spatio-temporal universe of natural entities studied by science is all there is. Global naturalists (e.g. Wilfrid Sellars) reject abstract objects of any kind, including traditional realist properties. Local naturalists (e.g. Jeffrey Poland) either accept or are indifferent towards abstract objects but they insist that the spatio-temporal universe consists only of entities studied by the natural sciences. Local naturalists reject Cartesian souls, Aristotelian entelechies, and so forth. A distinction also exists between strong and weak naturalists. Strong naturalists (e.g., David Papineau) accept a strict physicalism for the natural world, while weak naturalists (e.g. John Searle) embrace various emergent entities.

For someone to be a consistent naturalist, he or she must hold that the entity which emerges or comes-to-be (as naturalists understand that notion) (1) fits solely within a naturalist ontology (i.e. it bears a relevant resemblance to paradigm case natural entities of physics in its being and behavior); (2) can, at least in principle, be given a naturalist etiological explanation in keeping with the naturalist epistemic attitude and Grand Story if it is to be explained at all. The naturalist ontology must be selected so it will cohere well with the naturalist epistemology and Grand Story in order for the naturalist to be justified in asserting the scientific explanatory superiority of naturalism vs. rival world-views.

Elsewhere, I have argued that these aspects of naturalism imply that a consistent naturalist should be a strong naturalist.[38] I can only gesture at some salient points here. More than anyone else, D. M. Armstrong had clearly reflected on this topic and he concludes that all entities in the naturalist ontology must be (1) spatially located; (2) entities knowledge of which can be located within physical, causal processes; (3) capable of entering into causal relations and whose existence can be given a natural scientific causal explanation. To illustrate these points in Armstrong, the following statement is an example where he uses a naturalist externalist epistemology to settle issues in ontology:

> If any entities outside this [spatio-temporal] realm are postulated, but it is stipulated further that they have no manner of causal action upon the particulars in this realm, then there is no compelling reason to postulate them.[39]

In this context, Armstrong is claiming that the only way something can interact with natural entities – including cognitive processes so as to be objects of knowledge – is by way of causation.

Elsewhere, Armstrong employs the Grand Story as a criterion for an acceptable naturalist ontology: "I suppose that if the principles involved [in analyzing and explaining the origin of or processes of change in things within the single all-embracing spatio-temporal system which is reality] were completely different from the current principles of physics, in particular if they involved appeal to mental entities, such as purposes, we might then count the analysis as a falsification of Naturalism."[40]

Finally, as a naturalist, Armstrong explicitly rejects internal relations on the grounds that they cannot be spatio-temporally located.[41]

Strong naturalists eschew things like genuinely emergent mental properties/events for the same reason they set aside Cartesian dualism. Regarding Cartesianism, it is often claimed that dualist interaction is outside the constraints of naturalism because it requires an odd causal relation (i.e., one that relates entities from such different orders of being that the causal relation is no longer clearly a physical relation) and it postulates two radically different entities that enter into that relation one of which is not at home in the naturalist epistemology and etiology. Now these same points can and have been raised against a supervenience/emergence relation between physical and mental entities.

Perhaps enough has been said about contemporary naturalism to put us in a position to list without much development some of the main reasons why naturalists such as Sellars have embraced global and not local naturalism.

Naturalism and traditional realist properties

At least four reasons have been offered for the claim that global naturalism is preferable to local naturalism:

1. Traditional properties and spatio-temporal location: Some have offered a sort of "argument from queerness" against traditional realist properties to the effect that they are entities of a very strange sort, utterly different from anything else in the universe as depicted by the contemporary naturalist. Some have developed this argument by focusing on a special class of universals (e.g. propositions, axiological properties), others have claimed that no physicalist or naturalist non-reductive account of deep, metaphysical modality is possible and, thus, realist properties and their relations to each other are utterly

different from the rest of the naturalist ontology.

2. Traditional properties and the naturalist epistemology: As Laurence BonJour has recently noted, it is hard to accommodate traditional realist properties to physicalist depictions of human cognizers and of the causal interactions that constitute cognitive processes such as perception.[42] Given this depiction of humans and their various epistemic relations to things outside them, it is hard to see how we could have cognitive access to traditional realist properties. On a naturalist view, a necessary condition for knowledge is the ability of the knower to interact causally with the object of knowledge in standard physical ways. If properties are as traditional realists say they are, then they are causally inert abstract objects, yet we do seem to have such knowledge (or at least to know things that can be expressed using property-talk). Moreover, for those naturalists who eschew first philosophy and the conclusions reached therein, the entities associated with traditional realist properties (e.g. the properties themselves, the predication relation, the nature of an individuator) must be set aside. These entities are not "required" by our best scientific theories; they are allegedly justified if at all by considerations that are independent of and, arguably, epistemically prior to scientific concerns.
3. Traditional properties and being: Some argue that traditional realist properties violate the causal criterion of existence since they are non-spatio-temporal entities and remain that way even when exemplified. Moreover, given the non-spatio-temporal nature of properties, it becomes hard to see why their existence is dependent upon being exemplified. Thus, it is argued, uninstantiated universals exist and they clearly violate the causal criterion of being.
4. Traditional properties and predication: Traditional realism is a classic example of a "two world" ontology and, as such, it becomes difficult to explain in naturalist terms how there could be any kind of connection between abstract objects on the one hand and the spatio-temporal world of particulars and events on the other. Moreover, because the predication relation (i.e. nexus) (1) is non-spatio-temporal and (2) connects entities from different "worlds", it is hard to see how the relation itself bears a relevant similarity to strictly physical entities. In this regard, the predication relation presents the same sorts of problems to global naturalists that Cartesian interaction and emergent/supervenient relations do to weak naturalists (e.g. Where does Cartesian interaction take place and where is the predication relation exemplified?). The so-called explanatory gap between physical and mental events/properties has a parallel between spatio-temporal entities and those completely outside space and time. Finally, since at least in some sense, property-instances are ontologically dependent upon their properties, and since those properties are not themselves physical, if we take events as temporal states of affairs, the dependence of everything that exists or occurs upon physical factors in the physical domain becomes hard to maintain.

Rather than elaborate on these points, I want to turn to a consideration of naturalist strategies that accept these arguments and try to offer a naturalistically acceptable analysis of properties.

Two naturalist attempts to reconcile naturalism with acceptance of properties

Many naturalists, especially those sensitive to issues in first philosophy, have realized that (1) properties exist and (2) the traditional realist construal of properties is an uneasy fit with scientific naturalism. These naturalists also realize that traditional realism has a long and distinguished history and the onus is on them to develop an adequate theory of properties consistent with scientific naturalism. In my view, the two most powerful naturalist attempts at this project are those developed by Keith Campbell and D. M. Armstrong. In this section, we shall evaluate the adequacy of their efforts at such a rapprochement.

Keith Campbell

The most articulate version of nominalism currently available is the "trope" nominalism of Keith Campbell.[43] In this section, I shall describe the evolution of Campbell's nominalism over the last several years and criticize his current, refurbished version of nominalism. It will be instructive to analyze precisely the trajectory of Campbell's thought because such an analysis will reveal what is essential to a nominalist position. Part of the impetus for Campbell to change his views were criticisms that others and I raised against his earlier claims.[44] I hope to show that his refurbished nominalism is no more successful in avoiding certain problems than his earlier version. In what follows, I shall begin by investigating Campbell's earlier version of nominalism followed by an analysis of his most recent position.

1. CAMPBELL'S EARLY VERSION OF NOMINALISM

According to the early Campbell, the proper way to assay "Socrates is red" is this: the simple trope (Campbell's term for an abstract particular), red_1, which is a member of the similarity set "redness," is a part of the whole, Socrates, and Socrates is a bundle of compresent tropes.[45] This assay contains at least four key elements that made up Campbell's earlier position: the nature of a trope, the nature of a "universal," the relationship between a trope and its "universal," and the relation between a trope and a concrete particular that has it.

In Campbell's earlier writings, a trope was taken to be an abstract particular. By "particular" he meant an entity that is exhausted in one embodiment; by "abstract" Campbell meant an item which is got before the mind by an act of abstraction, that is, by concentrating attention on some, but not all, of what is presented. Examples of tropes are the specific, numerically singular taste of a lollipop, or its specific color, shape, etc.

Further, tropes are basic, primitive entities. Campbell says of a trope: "It embraces

no variety at all."[46] Again, "They are *infimae species*, taken as particulars."[47] In another place he says "They are in Hume's sense substances, and indeed resemble his impressions, conceived realistically rather than idealistically."[48]

On this view, a trope, construed as a basic particular, is simple, fundamental, and independent. By "simple" is meant that they have no further real parts (ones that can be separated off from the whole in reality and not merely in thought). The notion of being fundamental is explained this way. Entity A is fundamental relative to entity B if B's existence depends on A but not vice versa. Finally, "independence" means that an entity is truly independent if it could be the only entity in the universe. Tropes (at least non-emergent ones), then, are basic particulars. Concrete particulars like Socrates are derivative and it is a contingent fact that tropes usually come grouped together as concrete particulars.

An additional, important feature of tropes on Campbell's earlier view was the central role that space (or space-time) played in spelling out what a trope is. First of all, tropes do not exclude one another from being at the same place. In fact, a concrete particular is just a "clump" of compresent tropes at a place.

Second, a trope exists at one definite location; it is a quality-at-a-place. Campbell claimed that this should not be taken to imply that a trope has two different constituents, namely, its qualitative nature (which is a universal) and its place (which is particular). A quality-at-a-place is a single, particular reality. This is clarified by Campbell's view of the special status of the geometric figures of an entity, i.e. its form and volume. Form and volume are special tropes not like any others. Their presence in a particular compresent sum of tropes is not contingent; tropes cannot be present except by being in a formed volume. Tropes are, therefore, essentially regional. Wherever a trope is, there is formed volume. On the other hand, Campbell asserted that shape and size are not found except in company with characteristics. So formed volume is both essential to ordinary tropes and in itself insufficient to count as a proper being. Since the shape and size of a trope are inseparable from its color, then the former are not really different entities from the latter. Inseparable entities appear to be identical for Campbell; they differ only by a distinction of reason (when A and B differ by a mere distinction of reason, then A is identical to B). Put differently, the place (located, formed volume) and qualitative nature of a trope are identical and differ only as different ways of thinking of or speaking about the trope.

So much, then, for tropes. For Campbell, the relationship between Socrates and all the tropes "in" Socrates is a part/whole relation. Each trope is part of a bundle of compresent tropes which is Socrates. Thus the predication relation is reduced to a type of part/whole relation in Campbell's nominalism. The "universal" redness is merely a set of exactly resembling tropes, e.g. red1 (in Socrates), red_2 (in Plato), etc. In Campbell's earlier writings the relationship between each trope and its corresponding "universal" is clear and straightforward: it is the ε of set membership.

However, when we turn to an analysis of the exact similarity relation that obtains between or among tropes in a set, things are not so clear. In fact, the early Campbell's treatment of similarity is easily the weakest part of his version of nominalism. I have subjected Campbell's views on exact similarity to detailed analysis and critique elsewhere, and will content myself here with some general observations relevant to our current discussion.[49] One thing that is clear is that the unity of a set of tropes is grounded in the exact similarity relation. And the exact similarity relation is like the relation of identity in that it is symmetrical, transitive, and, arguably, reflexive.

On the other hand, it is not clear whether in Campbell's earlier views of exact similarity he saw it as an internal or external relation and he appears to have equivocated between the two views. To see this, we need to recall two things true of internal relations as they are usually construed, first, if the R of a to b is internal to a, then anything that does not stand in R to b is not identical to a. Second, internal relations are not primitive but, rather, are derived subsistences grounded in the natures of the entities they connect. As Gustav Bergmann put it,

> The ontological ground of an internal connection lies wholly 'in' the two or more entities it connects. More precisely, it lies in their natures. The notion is so crucial that I reword it. *The ontological ground of an internal connection is the natures of the entities it connects and nothing else.* Still differently, an internal connection has no ontological ground of its own . . .[50]

More recently, D. M. Armstrong has defined an internal relation as a relation which is logically determined by the nature of the related terms.[51] Armstrong goes on to point out that we can explain why internal relations are such that given two internally related entities a and b, there is no possible world in which the objects remain unaltered but in which the internal relation fails to obtain by recognizing that internal relations are derived from and grounded in the natures of the entities so related.[52] Since internal relations are ontologically grounded in the more basic metaphysical natures of the related entities, internal relations are derived subsistents, not primitive entities.

Now Campbell explicitly patterns his version of nominalism after that of D. C. Williams and Williams clearly held that exact similarity between two tropes is an internal relation.[53] Moreover, Campbell says,

> Can't we ask about two white tropes, no less than about two white things, what it is that gives them their resemblance? Perhaps we can answer legitimately just that they resemble each other because of what they are like, or because of their character or nature. Perhaps if we quantify over tropes we will have a cause to quantify over universals as well. But I for one have no confidence either way.[54]

In other words, in this passage Campbell does not clearly reject the internal relation view. But for three reasons, I think we should reject this interpretation. For one thing, in the passage just cited, Campbell grasps the fact that an admission that exact similarity is an internal relation commits one to universals and his entire system is meant to be an alternative to realism about properties. Second, elsewhere Campbell explicitly states that the relation of exact similarity is an unanalyzable, primitive, brute fact and this can only mean that the relation is an external one.[55] Finally, though Campbell claims that an anti-naturalist could embrace nominalism (e.g. by accepting the existence of mental tropes), nevertheless, he clearly wishes to advocate nominalism as a way of allowing the existence of qualities within the confines of a naturalist stance. And as D. M. Armstrong has pointed out, naturalists, if they are self reflective and consistent, should eschew the existence of internal relations.[56] I do not wish to defend Armstrong on this point, though I think he is right. Suffice it to say that if internal relations should be eschewed by a naturalist, then since Campbell wishes his ontology to be an expression of naturalism, he should avoid treating exact similarity as an internal relation. And it is charitable to interpret him in this way.

However, before I leave the topic of exact similarity, it should be pointed out that, at present, Campbell sees exact similarity as an internal relation.[57] It is important to note that Campbell's earlier confusion here, his change of position, and problems in his current treatment of exact similarity are rooted in a tension endemic to nominalism: (basic) tropes are simple entities – particularized natures – and there is a mere distinction of reason between either a trope's nature and location (or formed volume) or its nature and particularity. Thus, when one thinks about the fact that two red tropes are red, it is easy to see how their resemblance could be grounded in their color. But such a temptation does not arise if we think of two resembling tropes as merely located entities (or particulars). The relation of similarity between the "natures" of two red tropes is an internal one, but the relation between the location or particularity of those tropes is external. But how can this be if the nature and particularity of a trope differ by a distinction of reason and thus, are identical? I shall not probe this tension further here except simply to point out that Campbell's use of a distinction of reason here obscures his views and causes him to be inconsistent.

Recall that if A and B differ by a distinction of reason, then A is identical to B. With this in mind, I want to raise three puzzles that are problematic for Campbell's earlier views. Indeed, Campbell saw these puzzles in an earlier critique of mine and admitted that they were decisive against his position.[58] Consider a concrete particular, say, an apple. The taste trope and color trope of the apple, indeed all the apple's tropes, are at the same place since the apple is merely a bundle of compresent tropes. Now the nature of the taste trope differs from its location/formed volume by a distinction of reason only. Likewise, with the color trope and all the others. But then, the taste of the apple is identical to its place as are all the other tropes in the bundle. Now by the

transitivity of identity, all the tropes of the apple are identical to each other and, indeed, the apple is reduced to a bare location. Concrete and abstract particulars turn out to be bare simples and this is incoherent. This is puzzle one.

To understand puzzle two consider two compresent bundles of tropes, A and B. Bundle A includes a red and a taste trope at the same location. Bundle B includes a blue trope and a taste trope from the same exact similarity set as the taste trope in bundle A. In other words, the concrete particulars A and B differ in color but have the "same" taste. Now, the red trope of A is in a set of other red tropes exactly similar to it. Since the red trope's "redness" is identical to its place/formed volume (they differ merely by a distinction of reason) and since the nature of A's taste trope is also identical to the same place, A's red trope (by the transitivity of identity) is identical to A's taste trope. The same is true of B's taste and blue tropes. Furthermore, A's taste trope is in the same exact similarity set as B's taste trope. But this means that A's red trope and B's blue trope are in the same similarity set of taste tropes. And since exact similarity is transitive, B's blue trope will be in the similarity set of red tropes exactly like A's redness. If this is true, it is hard to make sense out of the idea of a similarity set based on a trope's nature. If an exact similarity set of red tropes includes a blue trope, the very idea of an exact similarity set of red tropes becomes vacuous.

The third puzzle is this. According to the early trope view, tropes are basic, independent existents and concrete particulars are derivative bundles of tropes. It is a purely contingent matter that a given trope is in a given bundle. The trope could have existed in another bundle of tropes or without being in any bundle at all. Now, a trope cannot exist if it is not individuated. But since place is the principle of individuation, this means that a given trope could not have existed at another place since that would make it a different trope. This same point holds for all the tropes at a given location which together form a compresent bundle. But then, it is not a contingent fact that these particular tropes combine to form a specific concrete particular, because in order that these tropes be *these* tropes, they must be at a given location. It is a matter of necessity that a concrete particular have just the tropes it does. In fact, the concrete particular is identical to the bundle of tropes, and, therefore, to each trope. And since the concrete particular is identical to each trope, it is identical to a mere place.

It is important to see just what is going on in these puzzles. The trope view must assay a basic trope as a simple in order to avoid assigning the individuating and qualitative roles to non-identical constituents in the quality-instance, for this is what realists do (e.g. red_1 has an individuator, say a bare particular expressed by 1, the universal redness, and a tie of predication). Since tropes are simples and since place/formed volume is identical to a trope's nature (differing by a distinction of reason only), then the trope nominalist must either remove the 1 and make the identity reflect the entity redness or else remove redness and make the identity reduce to bare location. The former collapses into realism and the latter is incoherent as the puzzles

indicate.[59]

2. CAMPBELL'S REVISED VERSION OF NOMINALISM

In light of these and other criticisms, Campbell has modified his trope nominalism.[60] Two aspects of his current thinking are especially relevant to our present concerns: his assay of a trope, especially the role of location in that assay, and his clarification of the relation of exact similarity. He continues to hold that a basic trope like red1 is a simple entity – a particularized nature that as a matter of brute, unanalyzable fact sustains two roles: to be a particularized nature that is exactly similar in nature to other tropes in its similarity set and that is individuated from all other tropes. A basic trope is not a union of a particularizing entity and a particularizer. It is a single, simple item – a particularized nature. Campbell is not entirely clear about the distinction between the nature and particularity of a specific trope.[61] On the one hand, he seems to continue to hold that the difference is just a distinction of reason since the nature and particularity of a trope do not constitute a genuine duality of entities which would turn tropes into complex entities. On the other hand, he says that the distinction between the nature and particularity of a basic trope is, perhaps, the formal distinction of Duns Scotus and, for Campbell, this amounts to the level of abstraction at which an entity is being considered (either as a particular or as a qualitative nature).

For two reasons I take it that Campbell's mention of the formal distinction really amounts to a reassertion of a distinction of reason in different terms. First, Suarez summarizes the medieval discussion of Scotus' formal distinction, correctly in my view, as being somewhat confused and hard to understand. But, nevertheless, based on the remarks of Scotus himself and his disciples, Suarez claims that the formal distinction turns out the be the distinction of reasoned reason.[62] Second, Campbell's own clarification of his use of the formal distinction reduces it to a distinction of reason since, for him, the distinction is an epistemological one only and identity obtains between the entity singled out differently.

Campbell goes on to say that a trope is not individuated by place. As a matter of a posteriori Kripkean necessity, tropes are located, but the location of a trope is a quasi-trope (roughly, a mere appearance or pseudo-entity that is part of the manifest world; more on this latter) and as a matter of necessity is always connected to a trope (it is contingent at which place a trope is located but not that it be located): ". . . let us abandon the view that a colour trope is individuated by its place. Take *compresence* as a more abstract, more formal matter, recognize that a colour trope and a spatial quasi-trope are distinct entities and assay the presence of green at a place as the compresence of a green trope with a place one."[63] I have not been able to tell what distinction Campbell sees between a trope's formed volume and its place. But regarding formed volume, he says that it is also a quasi-trope, a sub region of space whose boundaries are fixed by the presence of the trope's nature (e.g. color), that is

not identical to the qualitative content within that formed volume.

So much, then, for Campbell's new view of tropes. I have already mentioned that he has tried to clarify his view of exact similarity. In most places he says that this is an internal relation grounded in and supervenient upon the natures of the resembling tropes.[64] However, in other places he says that the exact similarity between two tropes is an ultimate, basic, unanalyzable, primitive fact.[65] There are at least two problems with these claims. First, in keeping with his naturalism, Campbell is committed to the idea that causal power is the mark of being.[66] I, for one, do not know what the ontological status is for supervenient entities in Campbell's ontology since they do not appear to have causal power. As we saw earlier, Jaegwon Kim has pointed out that for those with this metaphysical commitment, "To render mental events causally impotent is as good as banishing them from our ontology."[67] Kim's remark applies to Campbell's assay of exact similarity if that assay takes the relation to be a supervenient one. And Campbell himself admits that supervenient facts are mere pseudo-additions to our ontology in which no new being is involved.[68]

Second, if exact similarity is primitive, then no further analysis can be given either of the relation itself or of the fact that it obtains. But if the relation is internal, it is not primitive, not in the sense that the relation must now be reduced to identity – a reduction I accept but that Campbell is free to deny – but in the sense that exact similarity is grounded in the natures of the resembling tropes. As I hope to show below, once the need for a ground is admitted, difficulties follow that do not arise if such a need is denied. There are other interesting features of Campbell's nominalism, but rather than examine those features, it will suffice if we can see what is wrong with these two fundamental aspects of his, or for that matter, any version of nominalism.

My first set of remarks targets Campbell's new assay of a basic trope, starting with his claim that the new role that location or formed volume plays in his refurbished nominalism avoids problems that plagued his earlier position. Note first, that since a realist assay of a quality-instance treats it as a complex entity with its nature being a universal predicatively tied to an individuator, to avoid realism Campbell must hold that a basic trope is simple and this is where the main problems with nominalism lie. Apparently, Campbell fails to recognize that these problems turn on the simplicity of a trope and not on the role of location/formed volume in a trope's assay.

Moreover, Campbell's new view of the role of location/formed volume makes his nominalism less available to the naturalist than was his former position and Campbell is clearly concerned to present his trope nominalism as an attractive option for a naturalist ontology. Why is his new position more problematic for a naturalist? As we have seen, a widely accepted understanding of a naturalist ontological commitment is that the spatio-temporal universe of strictly physical entities that constitute the appropriate objects of natural scientific study is all there is. Campbell accepts this understanding of naturalism at least in the sense that he understands naturalism to

require all entities to be spatio-temporal. In his earlier view, it was easy to see why tropes could be viewed as natural entities that are entirely within the spatio-temporal world of natural entities. But the current picture of the role of location/formed volume actually comes perilously close to turning tropes into traditional abstract (non-spatio-temporal) entities, specifically, Platonic abstract perfect particulars with the proviso that they must be exemplified by a spatio-temporal particular to exist.

To see this, recall that for the traditional view of properties as abstract entities, properties are "in" the particulars that exemplify them, but they are not themselves spatio-temporally located at the place of those particulars. Nor is the way they are "in" those particulars a spatial relation; rather, it is the primitive relation of predication. Redness is (non-spatio-temporally, predicatively) in a ball and the ball is on the table, but neither redness nor predication are natural, spatio-temporal entities. It is widely recognized that when a universal is exemplified by a particular, the resulting state of affairs (the having of the universal by the particular) is itself a particular. This has been called the victory of particularity. Now in the same way there is a victory of spatio-temporality. When a ball is red, the ball is spatio-temporally located as is the state of affair consisting in the-having-of-redness-by-the-ball. But neither redness nor predication is spatio-temporal. If one were to add the Aristotelian constraint that properties must be exemplified in order to exist, the result would rule out transcendental (unexemplified) universals, but it would not turn predication or exemplified properties into spatio-temporal, natural entities. They would still be abstract objects.

Now Campbell's current views seem to parallel this picture precisely. A trope's nature is a metaphysically complete entity ontologically prior to its tie to the quasi-trope of location/formed volume since the nature is not identical to the location and neither enters into the being of the other in a part/whole way. This means that the trope's nature itself is not spatial but receives spatiality from its tie to the quasi-trope of location/formed volume. Nor is it clear that this tie is spatial since it connects a nature to a location. Now how is this different from the victory of spatio-temporality in the traditional realist view? The only difference is that a trope can be embodied only in one place. But not only is it difficult now to see why this is the case – if natures are identical to locations it is easy to see this, but since spatio-temporality is not inherent in their being, but rather, tied to them, it is less clear and more *ad hoc* to say such a tie can occur only once – it is also hard to see how a naturalist could appropriate such an analytic ontology into his or her world view. And even if we grant that a nature can be tied to space in one embodiment only, tropes turn out to be perfect particulars on one interpretation of Plato, and such entities were still non-natural abstract entities as was the tie that Plato claimed connects them to particulars in the space-time world.

Further, the simplicity of a basic trope in Campbell's refurbished nominalism generates the same metaphysical difficulties that he admits were telling against his earlier position. Consider two red tropes. Qua red entities, they stand to each other in

an internal exact similarity relation grounded in their natures. But qua particulars, they are spatially related to each other in external primitive spatial relation. But the particularity and nature of a trope are identical to each other (they differ by a mere distinction of reason). It follows from this that the two simple entities stand to each other in internal and external relations due to the same metaphysical features of the relata. But how can this be? Further, by the transitivity of identity, the particularities of the two tropes stand in an internal exact similarity relation to each other and the two red natures are externally related to each other. But this is just incoherent. Moreover, since the nature and particularity of a trope are identical to each other, then the same dilemma of simplicity presents itself to Campbell: either tropes are reduced to properties as universals or to bare particulars. This is a matter of ontology, not epistemology. And the distinction of reason, along with the epistemological use of abstraction, different modes of counting, and so forth obscure the real metaphysical problems involved here. Campbell's shift from using the distinction of reason between a trope's nature and location to using it between that nature and particularity does not solve the problem which centers primarily on the simplicity of tropes and not the special role of place in their assay. And as I have already pointed out, his new view of place is less, not more compatible with the naturalist stance.

Someone may respond that a similar problem can be raised against universals. A universal has a nature (what it is, e.g. redness for the universal red) and universality. If a trope's nature and particularity are identical for a trope, then the nature and universality of a universal are identical. Thus, red and green would be different by nature but identical by universality, which is incoherent. This response against the realist does not work. The realist need not take redness and greenness as simples. Therefore, if a realist grounds the universality of redness and greenness in a property possessed by both – being universal – he can distinguish this property from the natures of red and green. But the same option is not available to the trope nominalist, given the fact that a trope's nature and particularity differ only by a distinction of reason.

Here is an additional problem with the simplicity of a basic trope. A quality-instance, in this case a trope, must be extended in space or else it is not clear how a color nature can be present. A bare mathematical point of space cannot have a color nature at it. Unfortunately, something which is spread-out, i.e. extended, has to be a complex and not a simple entity because it has non-identical (perhaps, inseparable) parts at different spatial locations. In fact, it would seem that the very same nature (redness) occupies all the sub-regions throughout the extended formed volume. If an entity is to be extended (as an instance of color must be), it must be a complex entity (e.g. it contains non-identical inseparable parts) and a color trope, therefore, cannot be simple. G. Dawes Hicks raised this point against nominalist G. F. Stout when nominalism was just coming on the contemporary metaphysical scene earlier in this century:

> What, for example, in his view, is really meant by saying that 'qualities are in the same place as the things they qualify?' In what sense can it be asserted that a visual sense-datum *has* a specific colour? Each infinitesimal portion of such a sense-datum is *ex hypothesi* no less numerically distinct from every other portion than one sense-datum is numerically distinct from another. If, then, the colour of one sense-datum A cannot be identical with the colour of another sense-datum B, how can the colour of the part x of the sense-datum A be said to be identical with the colour of the part y of the sense-datum A? In other words, where, in such a case, do we ever reach a concrete particular thing which can be rightly said to have a particular character?[69]

Campbell is sensitive to these difficulties, especially to problems with stating identity conditions for spatially or temporally extended tropes that seem to constitute the primary and secondary qualities of ordinary macro-objects. Put briefly, his solution is this.[70] Campbell claims that basic tropes must be found that are partless, changeless, and unambiguous in their boundaries and he identifies five or six of them: space-time (construed as an absolute, non-relational trope), each of the four fundamental forces of physics taken as different space-filling fields that distribute some quantity across space-time to varying degrees, and, perhaps, consciousness understood as a pan-psychic field. By contrast, the manifest world is composed of the things and qualities, both primary and secondary, that are given to us in everyday experience. This world is composed of quasi-tropes: supervenient entities that are mere appearances. Such quasi-tropes are not truly real, and while they are not human inventions, nevertheless, there is a air of arbitrariness and conventionality about them regarding both their existence and identity conditions.[71] A manifest trope is a quasi-trope, a pseudo-entity, a mere appearance. Such quasi-tropes are taken in thought via abstraction to be separable entities. They are chunks of field tropes treated as if they were distinct and independent realities. For Campbell, an ontologist does not need to specify identity conditions for such multiply-located quasi-entities so my objections raised against them are wide of the mark. By contrast, the five or six omni-located basic tropes have clear identity conditions and that is all that should be required of Campbell's trope ontology.

Does Campbell's refurbished trope ontology provide an adequate response to the types of criticisms I have been raising? Is it an adequate analytic ontology and speculative cosmology? For at least three reasons, I think not. First, it is not clear that his response allows room for the types of considerations that justify the acceptance of tropist qualities in the first place. As far as ontology is concerned, the existence of such entities is justified in terms of various metaphysical or linguistic arguments from the manifest world, especially the phenomena of predication, resemblance, and abstract reference. To the degree that Campbell's basic tropes still exhibit these phenomena

(e.g. most of his tropes have the property of being a field), then to that degree he has not solved difficulties in regard to them but merely re-located those difficulties. If basic tropes do not exhibit these phenomena, then why accept the existence of basic tropes? A decision to do so seems to me to be an arbitrary choice. For example, Campbell calls space (or space-time) a trope. But he also tells us that it can change and expand. Space also has various properties, e.g. elasticity, extension, shape, volume, various other geometrical and topological properties, and the trait (whatever it is) that distinguishes empty space from the material stuff or fields that fill it. Given that Campbell is an absolutist regarding space, what is it that justifies treating it as a trope as opposed to a substance with various properties? He does not tell us and I suspect that, given the above characterization, most philosophers would take it in substantival terms.

Second, Campbell admits that the basic forces of nature are normally thought of as being exerted by various bodies with properties in virtue of which the fields of nature are generated.[72] On this view, the existence and properties of bodies are basic and fields are derivative. Campbell's ontology reverses this order. But so far as I can tell, he does not offer a single argument in favor of his preference. Perhaps this omission could be excused if his claims about these matters were merely speculative explorations of an ontology already justified on other grounds. But that is not the case. Campbell's field ontology is foundational to his trope nominalism precisely because his exposition of fields is meant as a response to serious criticisms that are both damaging to the entire trope ontology and used to justify a realist view of qualities. In light of this fact, Campbell's failure to justify his field ontology limits the acceptability of his view of tropes to those who agree with him about fields and, more importantly, leaves his analytic ontology resting on what amount to mere assertions in speculative cosmology. I doubt that this failure will convince many realists to adopt trope nominalism.

Third, Campbell's exposition of manifest tropes is so ambiguous that it is hard to tell what their ontological status is. On the one hand he says that the meaning of statements about supervenient entities cannot be captured by reductive paraphrases using only terms referring to subvenient entities, but for two reasons this seems to have no ontological implications for Campbell.[73] First, he explicitly says that supervenient entities are "pseudo-additions" and not real additions to our ontology; they provide cases that expand our descriptive resources without involving a commitment to the realities described. Second, for Campbell causal power is the mark of being and supervenient entities lack this power on his view.[74] The same ambiguity characterizes his statements about dependent entities like relations, manifest tropes, and sub-regions of space or fields. On the one hand, he says that relational facts exist but he also claims that they are non-existent supervenient entities. Such entities are not real because they are dependent entities. He also excepts foundationism, the view that all

relational facts can be accounted for in terms of non-relational foundational facts with no further ontology beyond the latter. So while talk about relations cannot be paraphrased away, relations do not exist.[75] Manifest tropes and other subregions of space or fields are quasi- or pseudo-entities that are not mere inventions, yet they are not distinct items and do not give us additional entities in our ontology.[76] They are not illusions, yet they are mere appearances that are treated as if they are independent entities by an act of abstraction and they do not really exist.

In my view, this confusion is the result of an inadequate set of distinctions. In the language of Francis Suarez, Campbell has no modal distinction that allows him to affirm the existence of dependent entities, e.g. Husserlian moments.[77] For Suarez, a modal distinction intervenes between an entity and its mode. A mode is a dependent, inseparable, genuinely distinct entity from that of which it is a mode. If a modal distinction obtains between two entities A and B (where B is a mode), there is a non-identity between A and B as well as inseparability in this sense: A can exist without B but not vice versa. For Campbell, only fundamental, separable entities exist. Talk about dependent entities may be irreducible but this has no ontological implications. Two inseparable "entities" differ by a mere distinction of reason. If I am right about this, than Campbell's ontology amounts to a denial of the existence of macro-objects and their qualities which I and, I suspect, others will find implausible. It may be that he accepts some sort of modes-of-being view in which dependent entities like manifest tropes are appearances in the sense that they have some sort of existence between non-being and full being. Now the difficulties with modes of being views of existence are widely recognized and I cannot discuss them here. Suffice it to say that if this is Campbell's position, he has not made that clear. But in any case, whether he denies or accepts a diminished sort of reality for manifest tropes, he would still face the difficulties I have been raising in the last few paragraphs. If manifest tropes don't exist, his basic ones still exhibit predication, resemblance, and abstract reference and the debate about nominalism will be located there. If they have enough existence to be colored, extended, and so forth, then problems of identity conditions for manifest tropes will still be relevant.

The simplicity of Campbell's revised tropes presents another difficulty for him: Campbell is inconsistent and, in fact, equivocal in his description of tropes. The fact that a trope is a simple entity that must simultaneously play two roles (have a qualitative nature that stands in the internal relation of exact similarity with other tropes like it and individuate) in a basic, primitive way is not a clear or natural way of thinking about quality-instances, and one must try hard to keep the view in one's mind without unwittingly treating the quality-instance as would a realist. This is why I think Campbell's own description of his position moves back and forth between what are apparently two inconsistent positions. I will mention two areas where this happens.

First, he repeatedly tells us that tropes *have* their natures.[78] On the other hand, he

also says that tropes are identical to their natures.[79] Which is it? The having relation is not the same as the identity relation (the latter is reflexive and symmetrical, the former is not) and it is the simplicity tension that causes this lack of clarity. Second, he says that the exact similarity between two tropes is an internal one grounded in the trope's natures and he also claims that this relation is primitive and unanalyzable.[80] The former view tends to treat tropes as complex entities that have natures as constituents and the latter tends to view tropes as simples. Of course, the realist will accept the former horn and claim that quality-instances *have* natures that ground the derived relation of similarity between two instances with the same nature. Campbell clearly needs the second horn – two tropes of the "same" kind are simples which stand in a primitive similarity relation – and it is this option that suffers from the problems I have been raising.

Campbell is aware of these problems about simplicity and so far as I can tell, he has two basic responses to them. First, he offers a sort of *tu quoque* argument against me and others who hold to bare particulars as individuators. He says that even though tropes qua simples must simultaneously sustain two different roles, the same problem attaches to bare particulars: they are simple and must simultaneously individuate and particularize.[81] According to Campbell, a particularizer gives the object in which it occurs a particular reality, as *a* definite object. But a specific object is not just an object, but this very object. And therefore, an object needs an individuator in addition to a particularizer, and bare particulars must sustain both roles.

Unfortunately, there is no genuine metaphysical distinction between particularity and individuation in this context. In fact, Campbell's treatment of them closely parallels Thomas Aquinas' discussion of undesignated and designated matter which, as Aquinas pointed out, differ by a mere distinction of reason with undesignated matter (e.g. flesh and bone in general) simply being a more general, abstract way of attending to designated matter (e.g. this flesh and bone). In Campbell's sense, particularity is not a different metaphysical entity from individuation. The former differs from the latter only in the level of epistemic abstraction one uses in attending to an object. Thus, there are not two metaphysical facts that need grounding as there are with so-called particularized natures (e.g. red and 1). If Campbell thinks these are, indeed, different metaphysical facts, then he has not given sufficient arguments for believing this to be the case. As far as I can tell, Campbell merely asserts his views in this regard without arguing for them. By contrast, I have offered reasons for thinking that a trope's nature and particularity are different metaphysical facts. Campbell admits that my arguments were successful against his identification of a trope's nature with its location/formed volume, but as I have shown, these same arguments apply to the particularity and nature of a trope.

Second, Campbell uses an illustration to argue that there is no problem with cases where a simple entity plays two roles: "A point is as simple as you can get. Yet points are

both different from all other points and also, for example, 10 mm distant from some others. That x is involved in more than one sort of fact does not show that x is complex."[82] But this example reveals a deep misunderstanding of the problem with tropes. A point is an individuated entity and this is so either because it has an individuator in it or else because it is identical to a bare individual. Either way, its particularity is a feature intrinsic to its being. But a point's spatial relation to another point is an external relation and is not intrinsic to the point. If that point were the only entity to exist, it would be an individual point but it would not be spatially related to anything else. Moreover, if a point is to the left of x but to the right of y, these two facts are not grounded in two different, intrinsic features of the point itself, but in two relations external to the point. By contrast, red_1 is intrinsically individual and red. If it were the only entity in existence it would be both red and particular. Thus, the ground for these two facts about red1 must be sought in its intrinsic being and not in external relations it sustains to other things. It is because red_1 is red that it exactly resembles other red tropes and it is because it is a particular that it is individuated from other entities. So Campbell's analogy breaks down.

We have seen that the role of location or formed volume in Campbell's nominalism, along with the simplicity of tropes, are serious difficulties in his understanding of qualities. A third problem area for Campbell involves abstract reference. Consider again statements like "red is a color" or "necessarily, red is a color." Now it is generally recognized that arguments based on inadequate paraphrases of such statements have been decisive against extreme nominalism. To see why, consider the scattered location, L, of all red things. An extreme nominalist might paraphrase the statements just mentioned as follows: "red things are colored things" or "necessarily, red things are colored things." Unfortunately, this paraphrase will not work, because red things are extended things or L located things (and this is necessarily the case for naturalist advocates of extreme nominalism), but redness is not extension or L location. The problem is that concrete particulars are complex entities in a way not true of first and second order universals (redness, coloredness).

In my view, the phenomenon of abstract reference is equally telling against nominalism. A nominalist would paraphrase the relevant sentences as "reds are colors" or "necessarily, reds are colors." But this will not work. In Campbell's earlier view, the counter example could read (with or without the modal operator) "reds are formed volumes" or "reds are L located" (where L is the scattered location of all red tropes). But redness is not formed volume or L location. In Campbell's current position, the counter example could read "reds are particulars," but redness is not particularity. Again, the problem here is not essentially one about location but complexity: in spite of what Campbell tells us, tropes exhibit a complexity (a nature and individuation) that is not like what is the case for first and second order universals.

Campbell's response to a closely related problem provides an insight as to how he

would handle this difficulty.[83] Campbell's strategy involves what D. C. Williams called painless realism. Consider a case where an interior decorator says he or she used four colors to decorate a house. Obviously, four tropes is not what is meant by this assertion and, in fact, the statement seems to view colors as universals. Nominalists, says Campbell, must make sense of cases where we appear to assert that matching tropes have something in common when, according to nominalism that cannot be the case.

Following Williams, Campbell makes a distinction between inherent and adherent properties. Inherent properties are the qualitative, monadic characteristics (e.g. being red) intrinsic to a specific item, i.e. that a complex concrete item can have in its own right. Adherent or "external" characteristics are relational properties of an object (e.g. being the heaviest man in the room). Now particulars can be "alike" in inherent (matching) characteristics and they can be distinguished by different adherent features, e.g. their compresent relations. The difference between a case (trope) and a kind (a "universal") is not a difference in ontological category, but rather, a difference in the rule used for counting. If we look at six red tropes which match in inherent but not adherent features, if you count one for each case, you treat the tropes as particulars. When, without regard to adherent characters, you only count more than one where there is a difference in inherent traits you take note of, you *identify* all occurrences of matching tropes, i.e. you treat them in a way that likens them to a realist universal. Applied to abstract reference, reds can be viewed as colors or as particulars, depending on what rule you use for counting.

I think painless realism is smoke and mirrors. For one thing, it is simply a label for the problem to be solved and is not a solution to the problem. The issue is not that we can notice a set of entities in a natural class and either count them as one entity on the basis of their common nature or as many entities according to their individuality. That is just what at least one version of the problem of the one and the many has always been about. The real issue is what grounds this activity and Campbell has no answer to this question. Moreover, his appeal to different rules of counting is just another example of the inevitable strategy of nominalists of reducing metaphysical issues to epistemological ones (e.g. try as they may, nominalist attempts to treat "abstract" as a metaphysical notion turn out, on closer inspection, to treat it an epistemic one). Indeed, they have no alternative since, for them, tropes are simples that sustain two functions, and these two functions will differ by the epistemic distinction of reason. In fact, Campbell's (and William's) appropriation of epistemology to solve a metaphysical problem obfuscates an obvious difficulty with the appeal to adherent features. If the problem of individuation is basically an epistemological one, namely, the task of counting, then it may well be that external relations are what we use to *single out* distinct entities in order to count them. But if the problem is metaphysical, adherent, external relations like compresence will not do because such relations presuppose and therefore cannot constitute the individuation of their relata.

Elsewhere, I have raised other arguments against Campbell's earlier views and

some of my points are equally relevant to his current position.[84] Rather than survey those problems here, I will content myself with the conclusion that nominalism, as exemplified by its most articulate paradigm case, has been shown to have enough serious difficulties to warrant a rejection of it. In spite of its subtle nuances, Campbell's refurbished nominalism is no more preferable than the position he himself has abandoned. Moreover, it is a less plausible option for a naturalist than his earlier position.

D. M. Armstrong

Armstrong is clear that a naturalist must find unacceptable uninstantiated universals, those classes of universals that are not the natural properties and relations characterizing the causal, physical objects of physics, and a priori knowledge. I shall not examine his case for these claims here. On the other hand, he believes that all versions of nominalism are unacceptable and that a robust version of realism about universals can be integrated with naturalism and physicalism. His own form of "realism" is just such an attempted integration and his model centers on explicating a view of (1) the nature of universals themselves; (2) the predication "relation"; and (3) individuation consistent with naturalism.

Regarding universals, the only way Armstrong can sustain realism and deny that universals are outside space and time is to reject the axiom of localization: no entity whatsoever can exist at different spatial locations at once or at interrupted time intervals. Armstrong rejects this axiom and claims that universals have multiple location and exist wherever the particulars having them are located. Thus, universals are entities that are spatially located at the place of the particulars that have them and they are fully present at each exemplified location.

Armstrong's treatment of predication and individuation are closely related. Universality and particularity are, he says, inseparable aspects of all existence, they are neither reducible nor related to each other and, though distinct, their union is closer than a relation. The distinction between universality and particularity is similar to Duns Scotus' formal distinction and is like the way that shape and size are united in a particular.[85]

This leads him to distinguish a thick from a thin particular. A thick particular is a state of affairs, e.g. A's being F, and as such, it is a particular along with its properties. The particular "enfolds" its properties in the sense that they are spatially located where the thick particular is. In the statement "this is hot," the word "this" refers to a thick particular and says that hotness is among its properties. The thin particular is the particular considered in abstraction from all its properties. It is not a thing *per se*, but amounts to bare numerical difference or thisness, the individuating factor that makes the thick particular more than just a bundle of universals.

Armstrong's view of individuation needs to be elaborated a bit more before it can be clear just what he means by it. He rejects transcendental realism in favor of immanent realism. We will see shortly that he is confused about the real nature of this distinction, but as he understands it, transcendental realism has universals standing entirely apart from their instances and immanent realism depicts them as being in the being of their instances, in Armstrong's position, being at the very place of their instances and being "enfolded" by those instances. Now Armstrong wants to reject what he calls relational immanent realism, roughly, the view that a thick particular is a complex state of affairs composed of a universal, a bare particular or substratum, and a relational tie of predication. Armstrong's basic argument against relational immanent realism is the claim that it falls prey to Bradley's famous regress argument against relations: if a substratum requires a relation to be bound to a property, then the substratum will require a relation to be bound to the first relation, and so on. Because of this, Armstrong claims that what is needed is a non-relational immanent realism. And his account of thick and thin particulars is just such an account. Says Armstrong, ". . . this version of Immanent Realism [is one] which distinguishes the particularity from the properties of a particular, while denying that the two aspects are related."[86] In sum, particularity and universality differ by a formal distinction, they are indistinguishable though inseparable, and "they" are too intimately conjoined to be related.[87]

Is Armstrong's version of realism a viable way to embrace properties within the constraints of naturalism? I do not think so and to see why, I want to focus on certain issues in Armstrong's view of properties as universals and on issues in particularity and individuation. Let us begin with properties.

First, is it not at all clear that his understanding of properties is, even if acceptable on other grounds, compatible with naturalism. In my view, naturalists should accept the axiom of localization and I think most naturalists would agree with me in this regard. Paradigm case naturalist entities would appear to obey what I shall call the Principle of In toto Location (PIL):

PIL: If entity e exists in toto at location L, then both (1) if x has proper parts, they overlap with and only with sub-regions of L, and (2) neither e nor any of its proper parts overlap with some other location P such that P is not identical to L or a sub-region of L.

If a necessary condition for being an entity in a naturalist ontology is conformity to PIL, then Armstrong's universals are not natural entities. It is just not clear how entities that are fully, completely present at each embodiment are relevantly similar to other entities embraced by naturalism to yield a consistent, economical ontology.

Besides considerations of PIL, most naturalists would find it difficult to believe that an entity could simultaneously be in motion and at rest if such a view is even coherent.

But Armstrong's properties are precisely such entities. If stationary particular x has F and moving particular y has F, then since F itself is at the location of the particulars, F is both moving and at rest. In my opinion, his universals are just as unacceptable for naturalism as is their construal as abstract objects.

It may be objected that only particulars must obey the laws of motion and properties are under no such constraint. However, to his credit, Armstrong realizes that a consistent naturalist is most reasonably taken to require that universals resemble particulars in having spatio-temporal location. But, then, why not make the same requirement regarding the prohibition of being simultaneously in motion and at rest? If Armstrong relaxes the second requirement, why not relax the first one? At the very least, Armstrong owes us an account of this problem and to my knowledge, he has not offered one.

Second, it is not clear that properties are located at the place of their instances.[88] For one thing, many properties (being even, triangularity) and relations (the musical relations between sounds in a music chord, being the father of) are such that it is just not clear what it would mean to take them as spatial, natural properties or relations. Armstrong could reject such entities on the grounds that they do not have causal powers and, thus, they have no being. But I suspect that many philosophers would find it more reasonable to reject causality as the mark of being than to eschew these entities.

Moreover, if a red ball is on the table and Armstrong claims that redness is spatially located on the table where the ball is, a reductive paraphrase is available to an advocate of properties as abstract objects for such a situation: the ball is on the table and redness is "in" the ball so the ball is located but the redness is not. The realist paraphrase allows one to say everything that needs to be said as far as spatial location is concerned, and the paraphrase includes a spatial particular (the ball) and two non-spatial entities (redness, and the way it is "in" the ball). Further, the instance of redness (see below) may be spatial, but the redness "in" that instance need not be. Put differently, every spatial predicate relevant to the state of affairs in our purview can be ascribed to either the ball or the instance of red (Husserl called it a moment)[89]. But redness itself is not spatial. When one attends to something spatial, one is attending to an instance of redness (or the ball itself), and not to redness. Redness is a color, it has hue, intensity, saturation, it resembles orange more than it resembles blue, and so forth. All the facts we know about red itself are not spatial facts. Spatial predicates are true of instances of redness, not redness itself.

Third, is spite of what Armstrong claims, statements like (1) "red is a color" or (2) "red resembles orange more than it resembles blue" express necessary truths. Traditionally, these have been understood as truths of reason, synthetic a priori truths that are true throughout possible worlds, including those with no red particulars. In general, for some universal F, there is a possible world w_1 such that F exists in w_1 but

no objects exemplify F in w_1.[90] Propositions like (1) and (2) are true throughout possible worlds and they are grounded in a view of universals as necessary beings whose existence is independent of the existence of particulars exemplifying them, and in genus/species or internal relations among universals. For example, (1) expresses a genus/species relation between a determinable (being colored) and a determinate (being red). And (2) expresses internal relations that obtain in a quality order consisting of the various determinates of a higher order, determinable universal. And arguably, both the truth and modality of these propositions are knowable a priori.

As far as I can tell, Armstrong has two main arguments against such universals. As we have already seen, the first one turns on his causal criterion of being. I find this criterion both inadequate and question-begging in this context. The second one represents a serious misunderstanding of what it means to say that some universal F is a transcendental universal. Armstrong seems to think that an advocate of transcendental universals must take such entities as "standing apart" from the spatio-temporal world, i.e., properties stand apart and exist separated from their instances in a different realm. The properties do not enter into the being of their instances.[91] This leads Armstrong to embrace what he calls the Principle of Instantiation: for each property P, there exists (not necessarily now) a particular, x, such that Px.

Unfortunately, Armstrong's statements in this regard are confused. For one thing, the debate between Platonists and Aristotelians about the relationship between properties and their instances is not a debate about the "location" of those properties (e.g. either in Plato's heaven or spatially contained in their instances), but rather, about the ontological independence of properties from their instances. Second, I have already shown how someone who sees universals as transcendental, abstract entities can still take it to be the case that universals enter into the very being of their constituents. Armstrong seems to think that the reading of Plato which turns properties into perfect particulars that are merely copied in their instances, or something of the sort, is the only view available for the transcendental realist, but this is just mistaken. The issue here turns on the way a property is "in" its instance.

Third, Armstrong's statement of the principle of instantiation reveals what I believe is a problem with his position. He claims that a property can exist when is not currently being instantiated as long as it will be instantiated at some time. Presumably, this could allow him to embrace properties ontologically (not temporally) prior to the Big Bang or properties (e.g. certain colors) that have only recently been exemplified. But how can the current existence of some entity depend upon the fact that in the future it will be exemplified in space and time? And how are we to conceive of the current mode of existence of such an entity? How does it currently fit into a naturalist ontology? Moreover, even if we grant the principle of instantiation, it does not follow that properties are spatially located at the place of their instances, but only that their

existence depends on that of the instances themselves. It could still be the case that properties are abstract entities "in" their instances in a non-spatial way. I conclude, therefore, that Armstrong's view of transcendental universals is historically inaccurate, a caricature of the position itself, and involved in problems in its own right. For these reasons, I do not think he has made his case against the existence of transcendental universals.

When we turn from properties to his account of particularity and individuation, things get no better. I, for one, cannot understand what Armstrong means here. If he takes the distinction between universality and particularity to be a formal distinction between "entities" that do not need to be related, then this amounts to the claim that they are identical, as we saw in our discussion of Campbell's tropes. If, on the other hand, he says that the thin particular (entity a) is a constituent in a thick particular (entity A) that is non-identical to a property/properties of A, then his view seems to reduce to a form of immanent realism with bare particulars. In this case, a thin particular turns out to be a bare particular that individuates, and a thick particular is a particular in virtue of its possession of the thin particular. As we have seen, this is the victory of particularity. But we have also seen that the victory of particularity is paralleled by the victory of spatiality, so on this latter reading of Armstrong, it is easy to see the universal in A and the predication relation between that universal and the thin particular as non-spatio-temporal entities. If this is the most satisfactory analysis of a thick particular, then given Armstrong's own statements about what is required of a naturalist analysis of properties, it counts as a refutation of naturalism.

On the other hand, it may be that Armstrong's earlier views on particularity and individuation collapse into nominalism, and it is no accident that in his later writings, he has opened up considerably to nominalism compared to his earlier writings.[92] In one place he says, "So provided you abandon uninstantiated universals (good riddance, I say), and provided Universals theorists and Trope theorists coordinate their views on just what properties and relations the world contains, it is easy to pass back and forth between the theories."[93] Indeed. Armstrong's claims that universals are spatio-temporally located at the place of their instances and his inadequate view of individuation are virtually indistinguishable from the trope nominalism of Campbell which, as we have already seen, is riddled with problems.

However, the various problems with nominalism have pushed Armstrong towards bare substrata though a failure to make a distinction in two senses of "being bare" obscures Armstrong's views on this subject. Advocates of bare particulars distinguish two different senses of being bare along with two different ways something can have a property. In one sense, an entity is bare if and only if it has no properties in any sense. Now bare particulars are not bare in this sense. They do not exist unless they possess properties.

There is another sense of bare, however, that is true of bare particulars. To understand this, consider the way a classic Aristotelian substance has a property, say,

some dog Fido's being brown. On this view, Fido is a substance constituted by an essence which contains a diversity of capacities internal to, within the being of Fido as a substance. These capacities are potentialities to exemplify properties or to have parts that exemplify properties. The capacities are grounds for the properties like brownness that Fido comes to have. When a substance has a property, that property is "seated within" and, thus, an expression of the "inner nature" of the substance itself. Thus, Richard Connell is correct to distinguish the way substances and bare substrata have properties when he claims that properties are not simply tied to substances, but rather "rooted in…and caused by the substance."[94]

By contrast, bare particulars are simple and properties are linked or tied to them. This tie is asymmetrical in that some bare particular x has a property F and F is had by x. A bare particular is called "bare," not because it comes without properties, but in order to distinguish it from other particulars like substances and to distinguish the way it has a property (F is tied *to* x) from the way, say, a substance has a property (F is *rooted within* x). Since bare particulars are simples, there is no internal differentiation within them. When a property is exemplified by a bare particular, it is modified by being tied to that particular. Thus, bare particulars have a number of properties, e.g. being red, and they have some properties necessarily, e.g. particularity, in the sense that a bare particular can exist only if it has certain properties tied to it. Now, this fact about bare particulars neither makes them identical to their properties nor does it entail that properties are constituents within a bare particular. Just because a man never comes out of his house naked, it does not follow that he is his clothes or that they compose him as constituents.

Armstrong's failure to distinguish these two senses of "being bare" have lead to confusion in attempts to interpret his account of individuation because he explicitly rejects yet implicitly seems to accept bare particulars. He rejects bare particulars for the following reason: "A particular that existed outside states of affairs would not be clothed in any properties or relations. It may be called a *bare* particular. If the world is to be a world of states of affairs we must [reject] Bare Particulars."[95] Yet as we have seen earlier, Armstrong makes a distinction between a thick particular (a's being F) and a thin particular (a). He calls the latter a Lockean substratum and accepts it as an individuator. For Armstrong, the particularity of a normal (thick) particular is an irreducible, unanalyzable feature of normal particulars. For him, particulars qua particulars do not differ in nature, but in virtue of bare, numerical difference, bare particularity, and he explicitly identifies this individuator (the thin particular) with a Lockean substratum.[96] Armstrong rejects bare particulars in the first sense but not the second, and given the fact that advocates of bare particulars embrace the second sense, his view is a bare particular position.[97]

There is a lesson to be learned from Armstrong's developing position. If one accepts a realist construal of properties, then one must also embrace some type of

individuator that is not a normal property (e.g. an impure property) or is not a property at all, or else the position collapses into nominalism. In his most recent treatment of the subject, Armstrong has acknowledged this point.[98] He continues to embrace universals, he explicitly rejects the individuative adequacy of spatio-temporal location, tropes and impure properties, and he takes it that "bare numerical difference" is what grounds the particularity of thick particulars. But on this reading of Armstrong, it is hard to see how his account is superior to a traditional realist view of individuation and particularity. In fact, it is hard to see how his position avoids collapsing into the traditional realist account. If I am right about this, then, in the absence of a better realist attempt to coordinate a joint commitment to universals and naturalism, and in view of the difficulties with both extreme nominalism and nominalism, the traditional historical position on the matter seems to be the correct one: the existence and nature of properties refutes naturalism as a world-view.

At this juncture, let us recall Johanna Seibt's assertion that "since the basic relation of a Platonist theory of predication, i.e. exemplification, cannot be defined in naturalist terms, a nominalist theory of predication proves to be '. . . the very foundation of a naturalist ontology'."[99] Among other things, my critique of Campbell and Armstrong has focused on their analyses of predication. I have tried to show that their accounts fail to be adequate as naturalist theories. I have also claimed that they are intrinsically problematic in their own right. To reinforce this second claim, it may be helpful to compare my traditional realist account of properties, predication, and property-instances mentioned earlier with the views of Armstrong and Campbell. In my view, there are at least four advantages to my traditional realist assay of an property-instance when it is compared to views like those of Armstrong or Campbell.

First, a universal F-ness must be brought together with its instances in such a way that they can be called F. Redness must be "tied to" its instances so as to make them red and not bare entities (or green for that matter). My traditional realist assay of property-instances explains why red_1 is red and not bare (or green). Red_1 is red because the universal, redness, is in it as one of its constituents and redness is the color of red_1.

Second, my assay of red1 clarifies the relation between redness and red1 in a way that is not found in views that liken this relationship to a set membership relation (Campbell) or a "being wholly spatially located at the same place as" relation (Armstrong). These treatments of this relation leave unclear precisely how the "universal" is really related to its instances. By contrast, the traditional realist sketch presents the universal as a constituent *in* its instances and, thus, treats this relation along the lines of part/whole relations. This part/whole relation involves a more deeply grounded relation between redness and the individuator in red_1. Admittedly, this relation is basic and undefined. But it can be viewed as a tie or a link between two entities. Moreover, the response to the claim that this version of "immanent relational realism" falls prey to Bradley's regress argument is one that Armstrong himself admits

is hard to refute: it is a unique feature of relations (including the ones included in our assay of qualities and quality instances) that they can relate relata without needing other entities to relate them to those relata.

Third, my traditional realist account presents red_1, as a complex entity with two constituents that serve to ground both the universality and the particularity of red_1. By contrast, instances for Campbell and, possibly, Armstrong (if his formal distinction is merely a distinction of reason) are simple entities without further entities in them. In a way that should now be familiar, this move leaves the abstract particular as a universal (red), a bare particular (1), or some unclarified entity that somehow sustains both functions without grounding either.

Finally, my account explains how it is that the universal transcends its instances. The universal is indifferent to any particular instance (and as a Platonist, I would add *all* its instances) since the universal can be a constituent *in* many instances through the non-spatio-temporal, inhomogeneous nexus of exemplification. When redness has red_1, as one of its instances, this is due to the fact that some entity (a bare particular) outside the nature of redness, has entered into an exemplification relation with redness. Something happens to redness, namely, it is modified and becomes exemplified. Thus, redness transcends its instances. If a universal is just a set of tropes (Campbell) then the universal cannot transcend its instances because the identity conditions for sets are constituted by the members of those sets. Change the membership and you change the set. Armstrong's views are more clear than Campbell's in this regard because he can claim that a universal transcends a specific instance in that it is wholly present at the same location as its other instances. However, this solution goes against the axiom of localization and for those who accept that axiom, Armstrong's model of transcendence will be problematic.

I have tried to show that if properties exist and are what traditional realists take them to be, then a widely accepted contemporary version of naturalism is false. It would seem, then, that one must either accept the traditional realist view and abandon naturalism, try to reconcile the two, offer a more adequate revised view of properties consistent with naturalism, or accept naturalism and extreme nominalism. I have attempted to offer at least some reasons why I think the first alternative is the best one to take. Whether or not my case has been persuasive, one thing seems clear: the ontological status of properties will continue to be a central issue of debate for those who have taken the naturalistic turn.

Notes

1 It is, of course, always dangerous to claim that there is a single, traditional view of some perennial issue in philosophy and I am making no such claim. Clearly, there have been different nuances given to realist views of properties as universals in the history of

philosophy. Still, I think it is safe to say that most would identify the traditional realist view of properties in the minimal way I have characterized it.

2 Howard Robinson, *Matter and Sense* (Cambridge: Cambridge University Press, 1982), 50. For more on this, see J P. Moreland, "How To Be a Nominalist in Realist Clothing," *Grazer Philosophische Studien* 39 (1991): 75–101.

3 See Reinhardt Grossmann, *The Existence of the World* (London: Routledge, 1992), 1–45; cf. my review of Grossmann in *Mind* 102 (July 1993): 407–10.

4 Wilfrid Sellars, *Naturalism and Ontology* (Atascadero, Calif.: Ridgeview Pub. Co., 1979), 109.

5 Ibid., 47. Cf. Wilfrid Sellars, "Towards a Theory of Predication," in *How Things Are*, ed. by James Bogen, James E. McGuire (Dordrecht, Holland: D. Reidel, 1985), 285–322.

6 Colin McGinn, *Mental Content* (Oxford: Basil Blackwell, 1989), 13. Cf. Jeffrey Poland, *Physicalism* (Oxford: Clarendon, 1994), especially pp. 10–44, 226–32, 307–12.

7 For a helpful discussion of this dialectic with a good list of resources, including the contributions by Field and Quine, see Penelope Maddy, *Realism in Mathematics* (Oxford: Clarendon, 1990).

8 See Maddy, *Realism in Mathematics*.

9 These are nicely offered in D. M. Armstrong, *Universals & Scientific Realism Vol. I: Nominalism & Realism* (Cambridge: Cambridge University Press, 1978), chapters 2–6; Michael J. Loux, *Substance and Attribute* (Dordrecht, Holland: D. Reidel, 1978), chapters 2–4.

10 D. C. Williams, "On the Elements of Being: I," *Review of Metaphysics* 7 (September 1953): 3–18; Keith Campbell, *Metaphysics: An Introduction* (Encino, Calif: Dickenson Publishing Company, Inc., 1976), 206–19.

11 Charles Landesman, "Abstract Particulars," *Philosophy and Phenomenological Research* 33 (March 1976): 323–37.

12 Gustav Bergmann, *Realism* (Madison: University of Wisconsin Press, 1964), 25.

13 Nicholas Wolterstorff, *On Universals* (Chicago: The University of Chicago Press, 1970), 130–9.

14 Woltersdorff, *On Universals*, 133–4.

15 Gareth B. Matthews and S. Marc Cohen, "The One and the Many," *The Review of Metaphysics* 21 (June 1968): 630–55.

16 J. R. Jones, "What Do We Mean by an 'Instance'?," *Analysis* 11 (October 1950): 1–9.

17 Edmund Husserl, *Logical Investigations*, 2 vols., trans. J. N. Findlay (London: Routledge & Kegan Paul, 1970), especially volume I, 408–10; II, 427, 436, 440, 453, 467–8.

18 See Grossmann, *The Existence of the World*, 1–45, 51–62.

19 See J. P. Moreland, "Was Husserl a Nominalist?," *Philosophy and Phenomenological Research* 49 (June 1989): 661–74.

20 Johanna Seibt, *Properties as Processes: A Synoptic Study of Wilfrid Sellars' Nominalism* (Atascadero, Calif.: Ridgeview, 1990), 184.

21 Cf. Alex Rosenberg, "A Field Guide to Recent Species of Naturalism," *British Journal for the Philosophy of Science* 47 (1996): 1–29.

22 Wilfrid Sellars, *Science, Perception, and Reality* (London: Routledge & Kegan Paul, 1963), 173.

23 Steven J. Wagner, Richard Warner, *Naturalism: A Critical Appraisal* (Notre Dame: University of Notre Dame Press, 1993), 1.

24 Roy Bhaskar, *The Possibility of Naturalism* (New Jersey: Humanities Press, 1979), 3.

25 John Searle, *The Rediscovery of the Mind* (Cambridge, Mass.: MIT Press, 1992), 11.

26 David Papineau, *Philosophical Naturalism* (Oxford: Blackwell, 1993), 3.

27 Robert C. Stalnaker, *Inquiry* (Cambridge, Mass.: MIT Press, 1984), 6.

28 John F. Post, *Metaphysics* (New York.: Paragon House, 1991), 121.

29 We should make a distinction between strong and weak externalism here. Focusing on justification, according to strong externalism, no factors that contribute to justification are internal to the agent. According to weak externalism, at least some factors that contribute to justification must be external to the agent. I am using the term *externalism* to stand for strong externalism because the weak version is not an option for consistent naturalists for the reasons just mentioned.

30 Searle, *The Rediscovery of the Mind*, 83–93.

31 E. Mayr, *Populations, Species, and Evolution* (Cambridge, Mass.: Harvard University Press, 1970), 4.

32 David Hull, *The Metaphysics of Evolution* (Albany, N.Y.: State University of New York, 1989), 74–5.

33 Bruce Aune, *Metaphysics: The Elements* (Minneapolis: University of Minnesota Press, 1985), 35.

34 D. M. Armstrong, "Naturalism, Materialism, and First Philosophy," *Philosophia* 8 (1978): 263. Cf. *Universals and Scientific Realism Volume I: Nominalism & Realism* (Cambridge: Cambridge University Press, 1978), 126–35. Subsequently, Armstrong has modified and weakened this formulation of his criterion of being: Everything that exists makes a difference to the causal powers of something. See *A World of States of Affairs* (Cambridge: Cambridge University Press, 1997), 41–3.

35 Daniel Dennett, *Elbow Room* (Cambridge, Mass.: MIT Press, 1984), 76.

36 Keith Campbell, *Abstract Particulars* (Oxford: Basil Blackwell, 1990), 172.

37 Jaegwon Kim, "Mental Causation and Two Conceptions of Mental Properties," unpublished paper delivered at the American Philosophical Association Eastern Division Meeting, Atlanta, Georgia, December 27–30, 1993, p. 23.

38 See J. P. Moreland, "Searle's Biological Naturalism and the Argument from Consciousness," *Faith and Philosophy* 15 (January, 1998): 68–91; and Moreland, "Should a Naturalist Be a Supervenient Physicalist?," *Metaphilosophy* 29 (January/April 1998): 35–57.

39 D. M. Armstrong, *Universals and Scientific Realism Volume I: Nominalism & Realism* (Cambridge: Cambridge University Press, 1978), 130.

40 D. M. Armstrong, "Naturalism: Materialism and First Philosophy," *Philosophia* 8 (1978): 262.

41 D. M. Armstrong, "Can A Naturalist Believe in Universals?," in *Science in Reflection*, ed. by Edna Ullmann-Margalit (Boston: Kluwer Academic Publishers, 1988), 111–12; *Universals & Scientific Realism Volume II: A Theory of Universals* (Cambridge: Cambridge University Press, 1978), 84–8.

42 Laurence BonJour, *In Defense of Pure Reason* (Cambridge: Cambridge University Press, 1998), 153–86, especially 156–61.

43 Cf. Keith Campbell, *Abstract Particulars*.

44 See J. P. Moreland, "Keith Campbell and the Trope View of Predication," *Australasian Journal of Philosophy* 67 (December 1989): 379–93; cf. J. P. Moreland, "A Critique of Campbell's Refurbished Nominalism," *The Southern Journal of Philosophy* 35 (Summer 1997): 225–46.

45 Cf. Keith Campbell, "Abstract Particulars and the Philosophy of Mind," *Australasian Journal of Philosophy* 61 (June 1983): 129–41; "The Metaphysics of Abstract Particulars" in *Midwest Studies in Philosophy Volume VI: The Foundations of Analytic Philosophy* ed. by Peter A. French, Theodore E. Uehling, and Howard K. Wettstein (Minneapolis: The University of Minnesota Press, 1981), 477–88; and Campbell, *Metaphysics: An Introduction* (Encino, Ca.: Dickenson

Publishing Co., 1976), 206–19.

46 Campbell, *Metaphysics*, 213.

47 Campbell, "Abstract Particulars," 141.

48 Ibid., 130.

49 J. P. Moreland, *Universals, Qualities, and Quality-Instances* (Lanham, Md.: University Press of America, 1985), 8–11, 41–5, 109–33. It is possible to identify exact similarity as identical to the entire set of exactly resembling tropes, or to identify different "kinds" of exact similarity relations with different sets of exactly resembling tropes (e.g. red tropes, sweet tropes), or to view exact similarity as a genuinely relational entity that is either a dyadic or an n-adic relation holding among n tropes. I will only consider the dyadic view because it is the most likely one to ascribe to Campbell and, in any case, it is the most plausible.

50 Bergmann, *Realism*, 54.

51 Armstrong, *Universals and Scientific Realism*, 172.

52 D. M. Armstrong, *Universals: An Opinionated Introduction* (Boulder: Westview Press, 1989), 43–4, 55, 100.

53 See D. C. Williams, "On the Elements of Being: I," *Review of Metaphysics* 7 (September 1953): 3–18; "On the Elements of Being: II," *Review of Metaphysics* 7 (December 1953): 171–92; "Necessary Facts," *Review of Metaphysics* 16 (June 1963): 601–26.

54 Campbell, *Metaphysics*, 216–17.

55 Campbell, "The Metaphysics of Abstract Particulars," 484.

56 D. M. Armstrong is one of the few naturalists to see and note this point. See, "Can A Naturalist Believe in Universals?" in *Science in Reflection*, ed. by Edna Ullmann-Margalit (Boston: Kluwer Academic Publishers, 1988), 111–12; *Universals & Scientific Realism Volume II: A Theory of Universals* (Cambridge: Cambridge University Press, 1978), 84–8. There are two main reasons naturalists must eliminate or reduce internal relations: (1) Such relations presuppose entities (e.g. essences, non-spatio-temporal quality orders like the colors, logical syllogisms taken as abstract objects) or a type of holism (e.g. living things are Aristotelian substances and not property-things) not consistent with naturalism. (2) It is hard to locate them spatially and temporally.

57 Campbell, *Abstract Particulars*, 37, 38, 59–60.

58 Cf. J. P. Moreland, *Universals, Qualities, and Quality-Instances*, 71–4; "Keith Campbell and the Trope View of Predication," *Australasian Journal of Philosophy* 67 (December 1989): 379–93. For Campbell's admission of the force of these puzzles, see Keith Campbell, *Abstract Particulars*, 65–6.

59 It is important to note that the problems revealed by these puzzles have nothing essentially to do with the fact that Campbell held to a bundle theory of substance. If one held a trope version of a traditional, Aristotelian-type view of substance, with minor adjustments, these same puzzles would apply. Dogs as much as reds exactly resemble each other and are individuated. And on Campbell's earlier view, a dog's nature would differ from its location by a distinction of reason. Elsewhere, I have criticized what amounts to a trope ontology combined with a traditional substance view. See J. P. Moreland, "How to be a Realist in Nominalist Clothing."

60 The primary place where Campbell has expressed his new position is *Abstract Particulars*.

61 Ibid., 56–8; 89–90.

62 Francis Suarez, *On the Various Kind's of Distinctions*, disputation 7, 1:13–15 (pages 24–7 of the 1947 Marquette edition).

63 Campbell, *Abstract Particulars*, 68–9

64 Ibid., 37, 59–60

65 Ibid., 31, 37–40.

66 Ibid., 67, 150–1, 172.

67 Kim, "Mental Causation and Two Conceptions of Mental Properties," 22–3.

68 Campbell, *Abstract Particulars*, 37. See Armstrong, *Universals: An Opinionated Introduction*, 56, 100. For Armstrong, the supervenience relation is the "nothing over and above" relation such that supervenient entities do not add any real, new entities to one's ontology.

69 G. Dawes-Hicks, "Are the Characteristics of Particular things Universal or Particular?," (in symposium with G. E. Moore and G. F. Stout), *Aristotelian Society Supplementary Volume III* (1923): 125.

70 Campbell, *Abstract Particulars*, 135–55.

71 Campbell says that if an entity is conventional, then it isn't real or natural, yet he also says that manifest tropes are not figments even though their identity conditions are arbitrary and conventional. These and other confusions permeate Campbell's discussion of these issues. See *Abstract Particulars*, 137, 152.

72 Ibid., 146.

73 Ibid., 37, 100–1, 103, 110, 121–2.

74 Ibid., 67, 150–1, 172.

75 Ibid., 98, 100, 101, 103–4, 110, 121–2.

76 Ibid., 125, 132, 145, 148, 150, 152–5.

77 See J. P. Moreland, "Was Husserl a Nominalist?."

78 Campbell, *Abstract Particulars*, 20, 22, 27, 56, 60.

79 Ibid., 27, 89.

80 Ibid., 37–8, 59–60.

81 Ibid., 9–10, 68–71.

82 Ibid., 70.

83 Ibid., 43–5.

84 Cf. J. P. Moreland, "Was Husserl a Nominalist?"; "Keith Campbell and the Trope View of Predication"; "Nominalism and Abstract Reference," *American Philosophical Quarterly* 27 (October 1990): 325–34; "Issues and Options in Exemplification," *American Philosophical Quarterly* 33 (April, 1996): 133–47.

85 In one place, Armstrong expresses a mild inclination to reject the literal multiple location of properties in favor of the view that all universals are "in" states of affairs that are themselves spatio-temporally located in the cosmos and, in this sense, are part of what constitutes the cosmos. See *Universals: An Opinionated Introduction*, 75–7, 81–2, 98–9. For three reasons, I shall not discuss this further. First, it runs contrary to almost everything Armstrong says about spatio-temporal location as a criterion of ontological commitment, e.g. that internal relations should be rejected because they cannot be located. This may be why he expresses himself here with the qualification that he is simply inclined to the view, not that it actually represents his considered judgment about the matter. Subsequently, he seems to have reaffirmed the spatio-temporal location of every existent. See *A World of States of Affairs*, 8, 39–41. Second, Reinhardt Grossman has pointed out that when Armstrong claims that all relations must be located in space and time except spatio-temporal ones, this concedes what the anti-naturalist argues, namely, that they are abstract objects, and Armstrong is clearly against the existence of the latter. See Reinhardt Grossmann, *The Existence of the World*, 27–8.

But this runs contrary to Armstrong's entire ontology. Finally, one of Armstrong's main concerns with uninstantiated universals is that if they exist, they exist "in another realm" outside space and time. Now, if the concern is to avoid such entities, it is not sufficient to claim that universals must be instantiated, because the traditional realist will say that properties remain abstract objects even while they are non-spatio-temporally "in" their instances. Thus, Armstrong needs to locate properties to avoid treating them as abstract; it is not sufficient merely to require of them that they be exemplified.

86 Armstrong, *Nominalism and Realism*, 109.

87 In his more recent thinking, Armstrong has followed Frege and sought to incorporate exemplication into the universal itself. See *A World of States of Affairs*, 28–31. He claims that properties are unsaturated entities that are filled by particulars, rather than being complete entities in their own right that are exemplified by particulars. According to Armstrong, this move allows him to avoid a Bradley-type regress that comes from admitting that exemplification is a relation and it makes sense of the fact that universals are dependent upon states of affairs and that there are no uninstantiated universals. Unfortunately, this move seems merely to be a verbal shift, for now universals must sustain two functions – having an intrinsic qualitative nature and being had by something – instead of having two entities each of which sustain one of these functions. It is hard to see how the former is preferable to the latter. Moreover, being unsaturated must now be taken as a real feature of universals different from their qualitative natures (e.g. since redness and greeness differ regarding their qualitative natures but agree in being saturateable). If so, it is hard to see how being unsaturated is a spatially located feature of either the universal or the state of affairs into which it enters. It is important to remember that for Frege, neither concepts nor their being unsaturated were spatio-temporally located. Third, there are other ways to avoid Bradley-type regresses without appealing to the notion of saturation and these strategies are well known. Finally, no gain is made for Armstrong regarding the dependence of universals on states of affairs or in providing a rationale for no uninstantiated universals. Frege himself believed that concepts could exist without entering a state of affairs (Armstrong's language) or without being saturated. It is just not clear why something could not exist while unsaturated. Without further argumentation which Armstrong fails to provide, a mere appeal to saturation and a particular's filling a universal provides no justification whatever for what Armstrong is after. Indeed, Frege's ontology, including his views about saturation, have long been seen as contrary to the naturalist stance.

88 Reinhardt Grossmann has shown that Armstrong's view of relations as spatial is even more problematic than is his position on properties. See Reinhardt Grossmann, *The Existence of the World*, 26–8; cf. 51–7.

89 See J. P. Moreland, "Was Husserl a Nominalist?"

90 See Michael Loux, *Substance and Attribute*, 92–6.

91 Armstrong, *Nominalism & Realism*, 64–6, 68–9, 102; *A Theory of Universals*, 3, 76.

92 See D. M. Armstrong, *Universals: An Opinionated Introduction*, 113–33.

93 Ibid., 122.

94 Richard Connell, *Substance and Modern Science* (Notre Dame: University of Notre Dame Press, 1988), 90.

95 Armstrong, *Universals: An Opinionated Introduction*, 94.

96 Armstrong, *A World of States of Affairs*, 124. Cf. 95–112; 123–7. Armstrong's only problem with Locke's view of a substratum, is that it was an unknown postulate and Armstrong insists that

the particularity of a thick particular is presented in acts of perception. Unfortunately, he is not as clear on this as one might wish because he seems to go back and forth between identifying what is so presented as (1) the fact that thick or propertied particulars are not exhausted by their properties and relations and (2) the thin particular itself. But as we have seen, the victory of particularity (which Armstrong accepts) does not clearly allow one to assert (2) just because one is acquainted in acts of perception with the thick particular qua particular.

97 For Armstrong's rejection of bare particulars in sense one, see *A World of States of Affairs*, 86, 153, 267–8. For a defense of bare particulars as individuators, see J. P. Moreland, "Theories of Individuation: A Reconsideration of Bare Particulars," *Pacific Philosophical Quarterly* 79 (September 1998): 251–63.

98 Armstrong, *A World of States of Affairs*, 107–11, 123–6.

99 Seibt, *Properties as Processes: A Synoptic Study of Wilfrid Sellars' Nominalism*, 184.

5 Naturalism and material objects[1]

Michael Rea

In recent decades, philosophers have devoted a great deal of attention to questions about how philosophical naturalists should understand mental objects, such as beliefs and desires, and apparently abstract objects, such as properties, numbers and sets. What has not been so often discussed, however, is the question of how naturalists should think about the wide variety of ordinary material objects that populate the ontology of common sense. This is the question that I shall address in the present essay.

At the outset, we must observe that our question is vexed – vexed by the fact that the house of naturalism is a house divided. There is little agreement about what naturalism is, or about what sort of ontology it requires. Some philosophers cast naturalism as a metaphysical view.[2] Others cast it as an epistemological or methodological view.[3] Still others deny that it is a single unified position at all, holding instead that there are several different, loosely connected *versions* of naturalism (metaphysical, epistemological, methodological, and so on).[4] Some naturalists allow abstract objects into their ontology;[5] and some say that naturalism is even compatible with belief in God, ghosts, souls, objective values, and other sorts of things widely held to be non-natural or supernatural.[6] Some say that naturalism, or some version thereof, implies or is even to be identified with either materialism or physicalism;[7] others deny that it is to be identified with any specific ontology at all.[8] Most naturalists would affirm Wilfrid Sellars's slogan that "science is the measure of all things: of what is that it is and of what is not that it is not";[9] and many no doubt would say that this slogan captures the heart and soul of naturalism. But a precise and even modestly uncontroversial explanation of what that slogan means has yet to appear in the literature. Hence, in one sense there really is no clear answer to the question of how naturalists should think about ordinary material objects because (as yet) there is no clear answer to the question of what it means to be a naturalist.

My own view is that naturalism is not a philosophical position at all, but rather a research program – a plan to conduct inquiry in a particular way. The naturalistic

program consists primarily of a plan to use the methods of the natural sciences, and those methods alone, in the development of philosophical theories. The consequences or implications of naturalism are just those views to which one is committed by virtue of following this program. Elsewhere I have argued at length that this way of characterizing naturalism is more faithful to the tradition than any of the other characterizations mentioned above, and it has the further advantage of avoiding such nasty problems as self-defeat and incompatibility with naturalistic methodological preferences.[10] Furthermore, the research program that I have identified with naturalism is one that seems to be universally endorsed by those who call themselves naturalists (one good reason for making the identification!), even if they disagree with the idea that acceptance of the research program is constitutive of their naturalism. Thus, by taking naturalism in this way, I am fairly well assured of defending conclusions that will apply to all naturalists, whatever their particular disagreements with one another might happen to be.

My thesis is that material objects as they are ordinarily conceived have no place whatsoever in the strict ontology of a consistent naturalist. On the ordinary conception of such things, every material object has the following property: facts about what sorts of changes it can and cannot survive do not depend upon human beliefs, attitudes, concepts, or conventions. More succinctly, but less precisely: material objects are *mind-independent*.[11] What I shall argue is that the methods of the natural sciences alone do not justify belief in mind-independent material objects; hence, naturalists have no business adopting an ontology that includes such things. I believe that in showing this, I am also showing that commitment to naturalism precludes belief in material objects entirely. This is because I believe it is impossible for there to be material objects that are not mind-independent in the sense just specified. But this claim is contentious, and I do not have the space to defend it here. Thus, I am officially defending only the more restrained conclusion.

The chapter has four parts. In the first, I argue that we can be justified in believing that there are mind-independent material objects only if we can be justified in believing that modal properties are exemplified in at least some of the regions of space-time that we take to be occupied by material objects. In the second, I argue that we can be justified in believing that modal properties are exemplified in a region only if we can be justified in classificatory judgments – judgments like 'this region contains an F', where 'F' is a name for a natural kind and furthermore constitutes a metaphysically better answer to the question, "What kind of object is in that region?" than any name for any other kind. In the third part, I dismiss as failures three views about how naturalists might be able to be justified in classificatory judgments. I consider a fourth which presupposes that, if there are material objects, then we are justified in believing that science reveals some of them to have proper functions; and I note that if this fourth proposal fails, there is good reason to think that any other proposal will as well. In the final section I argue that this fourth proposal does indeed

fail by showing that if there are material objects, science does *not* reveal any of them to have proper functions. If I am right, and if (as seems plausible) we are justified in believing that there exist mind-independent material objects (e.g. people), then we ought to reject naturalism.

I

Material objects have persistence conditions. By this I mean that for every material thing, there are facts about what changes it can and cannot survive. A man over six feet tall cannot survive being crushed down to the size and shape of a sugar cube, and a sugar cube cannot survive its own dissolution in a cup of tea. On the other hand, a man usually can survive the sorts of changes normally undergone in the course of an afternoon stroll, and a sugar cube usually can survive the sorts of changes normally involved in the daily existence of a sugar cube (being in a bowl, being jostled as nearby sugar cubes are removed, and so on). Even for instantaneous objects (if there are such things) which do not last over time at all, there are nevertheless facts about whether and in what circumstances it is *possible* for them to last over time. If there are no facts at all about what sorts of changes a putative thing X can and cannot survive, then there is no such thing as X.

Those who doubt this are invited to consider the claim that there is such a thing as Socrates but literally *no fact at all* about whether Socrates could survive a trip to Macedonia or a trip through a meat grinder. The claim is patently absurd; and the same can be said for similar claims about other material objects. Many are willing to accept *some* indeterminacy in persistence conditions. Thus, for example, one might think that if we were to annihilate Socrates' constituent atoms one by one, there would be no fact about *which* atom was the one whose annihilation finally did him in. But accepting this kind of indeterminacy is not at all the same as accepting the view that Socrates simply lacks persistence conditions. I know of no one who would say that there is no fact about whether Socrates could survive the simultaneous annihilation of *all* of his constituent atoms; and, again, similar remarks could be made with respect to any persisting material object. Where there are no persistence conditions, there is no object.

Having persistence conditions is part of our concept of a material object – just as having boundaries is part of our concept of a geometrical figure. But notice that there is a substantive ontological commitment lurking in the neighborhood. I have already indicated that persistence conditions are facts about what an object *can* and *cannot* survive; thus, they are modal facts (facts about what is necessary or possible for an object). Alternatively, they might just as well be construed as modal *properties* – properties like *being essentially such as to undergo mental activity* or *being essentially such as to have the chemical structure* H_2O (where X is essentially ϕ just in case X cannot exist

without exemplifying ϕ). Either way, acceptance of material objects carries with it a commitment to modality.

Once this is seen, it is perhaps easier to motivate the question: What justifies us in believing in material objects? At first blush, modal facts and modal properties seem wholly inaccessible to the natural sciences. Scientific methods can tell us, for example, that individual water molecules have the chemical structure H_2O; but it is hard to see how those same methods could tell us that water molecules have their chemical structure *essentially*, or that it is a necessary truth that water molecules have the chemical structure they do. Surely no experiment could confirm any of this. Moreover, the proposition that water molecules necessarily have the chemical structure they have doesn't obviously play any integral role in a total comprehensive scientific theory. We do not *have* to believe that water-molecule phenomena are caused by *material objects* in order to predict and explain those phenomena. Instead of an ontology of objects, we might simply adopt an ontology of stuff.[12] That is, we might suppose that our sensory experiences are caused not by discrete material objects but rather by variously propertied stuff, to which the notions of survival, generation, and destruction do not apply. There is nothing incoherent about supposing that nothing is destroyed when the stuff of which a water molecule is made is rearranged so as no longer to constitute a water molecule; and the same goes for other putative objects. In general, we do not need to believe in objects at all in order to make predictions and explain the various empirical phenomena associated with the (apparent) generation, existence and destruction of such things. Believing only in the stuff that constitutes them, along with its various qualitative properties, will do just as well. Thus, there is at least prima facie reason for worrying that, from a naturalistic perspective, belief in modal facts and modal properties – and so belief in material objects as well – will be unjustified.

One way of addressing this worry is to deny that there are modal properties and to maintain that modal facts are grounded somehow in human mental activity. To accept this sort of view is to embrace modal anti-realism. Modal anti-realism comes in two main varieties: conceptualism and conventionalism. According to conceptualism, modal facts are grounded in human concepts; according to conventionalism, they are grounded in human linguistic conventions.[13] The details of the two views are not important for our purposes. What is important is simply the fact that, with regard to saving the ontology of material objects, modal antirealism has one clear advantage and one fatal disadvantage. The advantage is that it makes it easy to see how we might acquire justified beliefs about persistence conditions: we simply reflect upon our concepts, or upon the empirically discovered qualitative features of the world and the semantic conventions governing our language. The disadvantage is that it is incompatible with the claim that material objects are mind-independent. Well, not quite: there are other possible versions of modal antirealism. For example, one might

think that modal facts are grounded in *divine* concepts or conventions or, less plausibly, that they are grounded in the concepts or conventions of other non-human beings (say, dolphins, apes, or Martians). Strictly, such versions of modal antirealism would be compatible with the mind-independence of material objects as I have defined it. But they are incompatible with naturalism, at least if we assume (as most naturalists do) that the methods of science alone provide no warrant for believing in other non-human beings with sufficiently rich concepts or languages to ground all of the modal facts in the world.[14]

If the ontology of mind-independent material objects is to be saved, then, refuge must be sought in modal realism.[15] For some, this fact will be sufficient to establish the conclusion of this chapter. Many philosophers, naturalists and non-naturalists alike, agree that naturalism is incompatible with modal realism. But many do not; and I do not need to establish anything so strong as the incompatibility of naturalism with modal realism in order to secure the conclusion of this chapter. The reason is that more than just modal realism is required to justify belief in material objects. Modal realism, but *not* belief in material objects, would be justified even if we were only justified in believing that sets or propositions have modal properties. To be justified in believing in material objects, we need to be justified in believing not only that there *are* modal properties but that such properties are exemplified in at least some of the regions that we take to be occupied by material objects. In the following sections, I will argue that commitment to naturalism precludes our being justified in believing that modal properties are so instantiated.

II

It is well known that essentialism about material objects – that is, the view that material objects have at least some essential properties – is plagued by the appearance of paradox. One particularly troublesome problem is what I have elsewhere called *the problem of material constitution*.[16] The problem is raised by a variety of apparently different puzzles, all of which reveal conflicting intuitions about what modal properties are instantiated in various regions of space-time. Consider, for example, a region of space-time occupied by a leafless oak tree.[17] Many people, at least initially, have the intuition that there is something in the region (namely, the tree) that cannot survive being chopped up into firewood, as well as something (namely, the aggregate of cellulose molecules that constitutes the tree) that can survive such chopping. The problem is that if the tree and the aggregate have different essential properties, it seems that they must be distinct; but if they are distinct, then there are two objects in exactly the same region of space-time, which seems impossible.

Reflection upon this puzzle reveals something about what is involved in our being justified in believing that particular modal properties are instantiated in particular

regions of spacetime. The puzzle arises because, in the region allegedly occupied by our oak tree, there exists matter that is arranged in at least two different ways – a way that is qualitatively distinctive of mere aggregates of cellulose molecules, and a way that is qualitatively distinctive of oak trees. The fact that the matter is arranged in these different ways is what allows us to set up the puzzle by saying that the region is both occupied by a tree and occupied by an aggregate. The puzzle is solved only when we figure out what the relationship is between having matter that is arranged in the distinctive qualitative way exemplified by members of a particular kind and having something that exemplifies the essential properties normally associated with that kind. But notice: we also need to be able to figure out this relationship in order to justifiably believe (apart from testimony) that particular modal properties are exemplified in a particular region. Solving the puzzle about the tree and the aggregate is precisely a matter of figuring out whether we have something that is essentially a tree, something that is essentially an aggregate, both, or neither, in the region allegedly occupied by the oak tree. In other words, it is a matter of figuring out which (if any) of a set of essential properties are instantiated in the region; and to do that, we must know something about the link between essential properties and ways of arranging matter.

Kind concepts are the tools by which we forge the link between essential properties and ways of arranging matter. It is traditionally assumed that there is both a metaphysical and an epistemological connection between kind-membership and persistence conditions. If this is right, then given the further assumption (which we shall continue to make throughout the remainder of this discussion) that persistence conditions are or are grounded in essential properties, there are also such connections between kind-membership and essential properties. But there are two ways of belonging to a kind. Something belongs to a kind in the *classificatory* way just in case terms referring to that kind are metaphysically better answers to the "What is it?" question for the object than terms referring to any other kind. On the other hand, something belongs to a kind in the *nominal* way just in case it exemplifies the distinctive qualitative features of those things that belong to the kind in the classificatory way.[18] On the traditional view, it is only by belonging to a kind in the classificatory way that an object is guaranteed to have the essential properties associated with that kind. Having one's matter arranged in such a way as to exemplify the distinctive qualitative features of the members of a kind is not sufficient. Thus, in order to justifiably believe that certain modal properties are instantiated in a region – indeed, to justifiably believe that any modal properties at all are instantiated in a region – we must somehow be able to distinguish between cases where there is *merely* matter arranged so that the distinctive qualitative features of the members of some kind are exemplified and cases where the matter really composes something that is classifiable as a member of that kind. In other words: we need to be able to distinguish regions that are nominally K-containing (for some kind K) from regions that contain something classifiable as a K.

(I should note here that I am assuming that every way of arranging matter is kind-constituting. But one might deny this, and in so doing one might come to the view that belonging to a kind in the nominal way entails belonging to it in the classificatory way. Thus, for example, Peter van Inwagen believes that the only genuine kinds are living-organism kinds;[19] and one might readily add to his view the claim that nothing can be nominally some kind of living organism without also being essentially that kind of organism. If this is right, then the problem we face in acquiring justified beliefs about the exemplification of modal properties is not one of distinguishing between regions that are nominally K-containing and regions that are essentially K-containing, for some kind K. Rather, it is a problem of distinguishing between ways of arranging matter that are kind-constituting and ways of arranging matter that are not. The two problems are equivalent, and they admit of equivalent solutions. Thus, I take it that my assumption that every way of arranging matter is kind-constituting is innocent.)

Are there other ways of acquiring justified beliefs about the essential properties exemplified in a region apart from testimony and apart from acquiring justified classificatory beliefs? Apparently not. Suppose you travel to a foreign planet and encounter a region containing matter arranged in a wholly unfamiliar way; and suppose a fellow traveler asks you whether there's anything in the region that could survive a blast from your laser gun. How could you even begin to answer? You could take a shot at it with your gun, but how would you know whether your shot destroyed anything? Even if you produced a gaping hole in the putative object, apart from testimony (e.g. squeals of pain from the object, or a booming voice imploring you not to take another shot) and apart from some justified view about *what the object is*, there would be no way of knowing whether you had damaged anything or simply rearranged some matter in an innocuous way. Thus, it seems that if we have justified beliefs at all about the exemplification of modal properties in regions of space-time, our having such justified beliefs depends either on testimony or on our having justified classificatory beliefs.

Obviously not *everyone's* beliefs about essential properties can come from testimony. Thus, our being justified in believing that particular essential properties are exemplified in particular regions of spacetime must ultimately depend upon someone's being justified in their classificatory beliefs, and this too apart from testimony. But, of course, commitment to naturalism implies that someone can be so justified only if the methods of science alone allow us to distinguish between nominally K-containing regions (for some kind K) and regions that contain something classifiable as a K. In the next section, I shall argue that the methods of science alone do not allow us to make such determinations.

III

I will begin by dismissing two suggestions about how we might acquire justified classificatory beliefs that are initially tempting but that are readily seen, upon brief reflection, to be non-starters. I will then consider what I take to be the two most promising explanations and show why they fail. I should say up front that, in order for my argument in this chapter to be strictly valid, I need the further premise that if the alternatives I am here considering fail to provide a naturalistically acceptable explanation for how we might acquire justified classificatory beliefs, then every other alternative explanation will fail as well. I cannot prove that this premise is true, since (as the reader can easily verify) the alternatives I am considering are not *logically* exhaustive. However, I see no reason to think that there are other even remotely promising alternatives; and so I think it is reasonable to assert on this basis that if these alternatives fail, any others will as well. Those who disagree may simply take my argument as a challenge to provide a successful alternative.

One might think that we can acquire at least one justified classificatory belief in the following way: we introspect and find experiential evidence for our own existence. We then infer from this that essential properties must be exemplified in at least one region of spacetime – namely, the region we ourselves occupy. Strictly speaking, this procedure saves belief in material objects. The trouble, however, is that it saves it only by presupposing it. We have experiential evidence of our own existence, but not of the fact that we are material objects. The belief that we are our bodies depends for its justification upon the belief that we have bodies – a belief which is justified only if we are justified in believing that the matter in the region putatively occupied by our bodies is more than nominally a human-body-containing region. The question is why we should think that it is, and so far we have no answer.

On the other hand, one might think that the belief that we are material objects is a Moorean fact – a datum of common sense, automatically justified and sufficiently so to constitute an argument against any set of premises that imply its denial. Though I think it is far from obviously true that materialism about the mind has the weight of common sense behind it in the same way that realism about the external world does, I need not press that point here. The main problem with this suggestion is that following it is not compatible with commitment to naturalism. If the belief that we are material objects is established through Moorean common sense *rather* than the empirical methods of science, then it is not justified from a naturalistic point of view.

So much, then, for non-starters. As I see it, there are only two even remotely promising views about how we might acquire justified beliefs about classification through the methods of science alone. The first may be treated rather briefly. Undermining the second will occupy the whole of Section IV.

According to the first of the remotely promising suggestions, our classificatory

beliefs are justified by their role in the formulation of natural laws. The belief that a particular region contains a tree rather than an aggregate of cellulose molecules, for example, might be justified on the grounds that similar classifications have provided a basis for the articulation of natural laws governing the behavior of trees which have then been confirmed through empirical testing. Other classificatory beliefs would be justified in similar ways. As a general rule, then, the contents of regions are to be classified in ways that allow for the simplest, most economical formulation of an overall set of natural laws; and the justification for the total classificatory scheme is precisely the same as the justification for the total theory of which it is a part.

This view suffers from two weaknesses. The first is that it makes the justification for all of our classificatory beliefs *pragmatic* rather than *epistemic*. The fact that our task of theory-construction is simplified by adopting one set of classificatory beliefs rather than another is not at all obviously a reason for thinking that they are true. The classificatory scheme we adopt, as well as the whole ontology of material objects itself, might well be a convenient fiction; but why should convenience warrant belief? I realize, of course, that many naturalists (though not all!) think that pragmatic justification of this sort is sufficient for epistemic justification (or – what might come to the same thing – that pragmatic justification is the best we could ever hope for in the way of justification). But even if I were to concede that pragmatic justification for our classificatory beliefs is sufficient enough for whatever justification we might desire for belief in material objects, there would still remain a further weakness with the view under consideration. That weakness is just that there seems to be no good reason to believe that our theories can be significantly simplified by adopting one set of classificatory beliefs rather than another. In biology, for example, there seems to be good reason to think that any of a variety of classificatory systems will do equally well; and in physics it seems that we can get along just fine by restricting our ontology to include only space-time, regions of space-time, and the various qualitative properties that might be exemplified at those regions.[20] Thus, even if pragmatic justification were sufficient to meet our needs, it is not at all clear that we could get it.

A more promising view about how we can acquire justified classificatory beliefs is (roughly) that we can acquire such beliefs by acquiring, through the methods of science alone, justified beliefs about proper function.[21] But here we must be careful. Obviously we cannot say that our classificatory beliefs depend for their justification upon our ability to acquire justified beliefs to the effect that some *mind-independent material object* has a proper function; for I have already argued that belief in mind-independent material objects depends for its justification upon our ability to acquire justified classificatory beliefs. Nevertheless, one might think that whether or not there *really* are mind-independent material objects in the world, still there are phenomena in nature revealed by scientific methods such that, *on the assumption* that there are mind-independent material objects, there is strong evidence for believing that science

reveals some of those objects to have proper functions. And one might think that it is these *proper-function-phenomena* that somehow ground our classificatory beliefs.

The view now under consideration may be expressed more precisely as follows. We can acquire justified classificatory beliefs through the methods of science alone if, and only if, (3.1) is true:

(3.1) If we were to assume that there are material objects, then there would be strong evidence for believing that science reveals some of those objects to have proper functions.

As I said before, the 'only if' is justified by the fact that no other alternative seems available. Those who doubt this are again invited to supply another alternative. In the remainder of this section, I will briefly explain why I think that (3.1) implies that we can acquire justified classificatory beliefs through the methods of science alone. In Section IV, I will argue that 3.1 is false by assuming that there are material objects and showing that there is no reason to believe that science reveals anything to have a proper function.

Before explaining why I think that (3.1) implies that we can acquire naturalistically respectable classificatory beliefs, I must make some preliminary remarks about function ascriptions. Almost everyone writing on functions agrees that function ascriptions are system-relative. In other words, statements of the form "x has the function of doing ϕ" are always elliptical for statements of the form "In system s, x has the function of doing ϕ." The same is true for ascriptions of proper functions. Nevertheless, we do sometimes speak of whole systems as functioning properly or improperly even when it is doubtful that the system as a whole has a proper function. For example, we speak of cars as functioning properly or improperly. We might even speak of ourselves, or our pets as functioning properly or improperly. (We *might* do this; but in many contexts it would sound strange. One who pointed at his rabid dog and said "Boy, it looks like my dog is malfunctioning!" would probably appear to be making a joke.) So *proper functioning* does not necessarily involve *performing a proper function*. In one legitimate sense of the term, a system as a whole is functioning properly just in case it is healthy (or, for artifacts, in a state relevantly analogous to health); and this is so whether or not the system itself *has a proper function*. Toward preserving these distinctions, then, I will say that something has a *proper function* or a *proper part-function* just in case it has a function it is supposed to perform within some larger system, and I will say that something has a *proper system-function* just in case there is some way of functioning that counts as proper functioning (or health or a relevant analog of health) for that system. That said, I turn now to my defense of the claim that (3.1) implies that we can acquire naturalistically acceptable classificatory beliefs.

In short, the idea is that the existence of proper-function-phenomena in nature might reasonably be construed as evidence to the effect that, though the matter in a

region is arranged in multiple ways at once, one of the ways in which it is arranged is privileged over the rest; and such evidence provides a basis for classification. Consider a region of space-time that contains matter organized in such a way that (a) there seems to be a system S in that region, and (b) S seems to have parts that have proper functions relative to S.[22] Suppose further that, on the assumption that there really are mind-independent material objects, the evidence is that the methods of science alone have revealed the proper part-functions of some of the parts of S. Then it seems clear that the methods of science alone have revealed that among the various ways in which the matter in the region allegedly occupied by S is arranged, one is metaphysically privileged over the rest as the way that matter is *supposed* to be arranged. But now we have a basis for a classificatory judgment, for, given that the matter in the region allegedly occupied by S is supposed to be arranged in a particular way, it is quite plausible to think that the region is occupied by an object classifiable as a member of whatever kind is associated with the particular way in which the matter in that region is supposed to be arranged. So if (3.1) is true then classificatory judgments (and hence modal attributions and belief in material objects generally) can be justified through the methods of science alone.

There is one very obvious objection that might be raised against the view just described. Even if the view does explain how we might be justified in our classificatory beliefs involving macroscopic objects like organisms and perhaps artifacts, it doesn't at all seem to explain how we might be justified in classificatory beliefs involving non-living natural objects like rocks, planets, water molecules, atoms, and the like. After all, neither planets nor water molecules nor any of their respective parts seem to have proper functions. Thus, at best, the view seems to provide the naturalist only with a way of admitting *some* material objects into her ontology. It does not offer the resources for countenancing the wide variety of non-living natural objects in which many of us believe.

There is a genuine concern here, but I do not think it is decisive. Suppose one thinks (with Aristotle, for example) that living organisms are *paradigmatic* examples of material objects, and that our beliefs about proper function do indeed ground our beliefs in those paradigmatic items. One might then go on to argue that our beliefs in other things (like rocks and atoms) are justified either by analogical extension, or by their role in scientific theories. I have already argued that the role of talk about such objects in scientific theories is not sufficient on its own to justify belief in their existence or in material objects generally. But it is quite plausible to think that once we have already become justified in believing in some material objects by way of our beliefs about proper function, there is no longer any obstacle to believing in the many other sorts of objects that are talked about in our scientific theories. Our beliefs about proper function give us reason to abandon a stuff-ontology in general, and once that is abandoned, it becomes legitimate to take all of our apparent talk about material objects at its face value.

Incidentally, my own view is that something like this story is the correct story about how we acquire justified classificatory beliefs (and hence justified beliefs in modal properties and material objects). That is, I am inclined to think that we in fact acquire such beliefs by way of our beliefs about proper function. However, as I shall argue in the next section, this sort of view about how we acquire justified classificatory beliefs is not available to the naturalist. The reason is that (3.1) is false: even if we assume that there are material objects, there is *no* good evidence for believing that the natural sciences reveal some of them to have proper functions.

IV

Let us begin the argument against (3.1) by observing that science can reveal the proper system-function of an object only if it can also reveal the proper part-functions of some of its parts. This is not necessarily to say that we learn facts about proper system-function by inferring them from facts about proper part-function (though perhaps we do). My point is simply that there is no reason to suppose that science can tell us what it is for an organism to enjoy proper system-function but leave us totally in the dark about what its various parts are supposed to do. Proper system-function supervenes on proper part-function; so if the former is accessible to science we would expect the latter to be also.[23] Furthermore, facts about proper system-function are not radically different in kind from facts about proper part-function; so, again, there is no reason to suppose that science can reveal the former but not the latter. (In the same way there is no reason to suppose that science can reveal facts about rights but not facts about obligations, or facts about mental phenomena but not facts about brain chemistry.) Thus, the naturalistic acceptability of beliefs about proper system-function stands or falls with the naturalistic acceptability of beliefs about proper part-function.

For my purposes, a part of a thing has a proper function just in case (a) it has a function, and (b) one of its functions is metaphysically privileged – i.e. the part is objectively *supposed* to have that function, or it absolutely *must* have the function, or its having that function contributes in some way to the Good (construed as objective and universal), etc. I think that these two conditions accord well with ordinary intuitions about what it is for something to have a proper function; but I should note that, whether they do or not, at any rate this is the sense of proper function that is needed in order for (3.1) to imply that we can acquire justified classificatory beliefs. If proper functions are not metaphysically privileged functions, then there is no reason to suppose that we can discover a metaphysically privileged arrangement of parts for a thing by discovering proper functions.

I will take it for granted that, on the assumption that there are material objects, there is good evidence for thinking that the methods of science alone can reveal that something has a function. One might doubt this. For example, some philosophers

think that an object has a function just in case it contributes to a goal of some system of which it is a part; and some philosophers think that a system can be goal-directed only if it has been designed. If both views are correct, then it follows that biological objects have functions only if there is a cosmic designer. But we are assuming that the methods of science alone cannot reveal the existence of a cosmic designer; hence, it is hard to see how (if this view were correct) the methods of science could reveal that anything has a function. Of course, if the methods of science *cannot* reveal that anything has a function, then *a fortiori* they cannot reveal that anything has a *proper function*. Hence, I take it that in assuming that science can reveal that some things have functions, I am simply giving the naturalist a running start.

There are various different views about what it is for something to have a function. Though it would certainly be interesting to do so, there is no need to canvass that territory here. All that is important for present purposes is the fact that, no matter what view of functions ultimately turns out to be correct, in the literature on functions there are only two suggestions compatible with commitment to naturalism as to what might make it the case that a function of a thing is also its *proper* function. One suggestion is that statistical normality is what determines proper function. For example, one might think that X has F as its proper function just in case X belongs to a kind whose members (statistically) normally have F as their function. The other suggestion is that the etiology, or causal history, of a thing determines its proper function. So, for example, one might hold that X has F as its proper function just in case X exists at all or exists where it is *because* it tends (or, better, because things of its type tend) to do F.[24] I will argue that, though science can indeed detect facts about statistical normality and etiology, neither statistical normality nor natural etiology is sufficient for proper function. Hence, lacking other naturalistically acceptable suggestions as to what properties might make it the case that a function of a thing is also its proper function, I will conclude that the methods of science alone cannot reveal that anything has a proper function.

Consider first accounts of proper function in terms of statistical normality (let's call them 'SN-accounts' for short). No one that I am aware of has explicitly proposed an SN-account. However, I discuss it because, as Alvin Plantinga notes, this sort of analysis is "widely popular in the oral tradition."[25] Plantinga himself offers several counter-examples; but most of them seem directed against SN-accounts of proper system-function rather than against SN-accounts of proper part-function. So, for example, he writes:

> Most 60–year old carpenters have lost a finger or thumb; it is not the case that those who have not have hands that are not normal and not capable of proper manual function. . . . Perhaps most male cats have been neutered; it hardly follows that those that haven't are abnormal and can't function properly. The vast

> majority of sperm don't manage to fertilize an egg; the lucky few that do can't properly be accused of failure to function properly, on the grounds that they do things not done by their colleagues. Most baby turtles never reach adulthood; those that do are not on that account dysfunctional. Obviously you can function properly even if you don't function the way most of your peers do.[26]

I take it that these examples do indeed show that statistical normality cannot be the whole story about proper system-function (and perhaps that is all they are intended to show). But none *directly* show that it is not the case that a thing X has F as its proper part-function just in case X belongs to a kind whose members statistically normally have F as their function. The reason is that none of Plantinga's examples seem to provide us with an X that satisfies the right-hand side of the biconditional. Hands, no doubt, have functions, but it is not at all obvious that *having a thumb* is among them. Sperm might also have various functions, but, again, we might well doubt that *fertilizing eggs* is among them. As I noted earlier, function ascriptions are system-relative; thus, to say that sperm have the function of fertilizing eggs is to say that there is some larger system of which the sperm are parts and in which they have the function of fertilizing eggs. But what system might that be? Certainly not the male organism, for sperm do not fertilize eggs in the male organism. But not the female organism either, for sperm are not *parts* of a female. Perhaps there is such a thing as Shakespeare's beast with two backs, and perhaps *that* counts as a system of which sperm are parts and relative to which they have the function of fertilizing eggs. But such a system is not a biological organism, and it is not obvious that there are biological systems other than organisms; hence it is not clear that fertilizing eggs can count as a biological function of sperm, and many will doubt that sperm, as biological objects, can have any functions but biological ones. And, of course, similar reasoning will strongly suggest that cats and baby turtles lack functions as well.

Still, Plantinga's examples do have indirect bearing on SN-accounts of proper part-function. They point out that there is no reason to think that what is statistically normal is metaphysically privileged in any way. Most male cats are neutered, but it does not follow from this that being neutered is a metaphysically privileged way of existing for a male cat, or that male cats are supposed to be neutered in any deep and interesting sense of "supposed to." Similarly, then, it is hard to see why an object's belonging to a kind whose members normally have F as their function should imply that it is, in some metaphysically important sense, *supposed* to have F as its function. The proper function of a human optic nerve is to transmit information to the visual cortex. But this is not because it is statistically normal that it do so. If it became statistically abnormal for optic nerves to transmit information to the visual cortex, then malfunction of the optic nerve would be statistically normal. It would not be the case that optic nerves were no longer supposed to play the role in vision that they in fact

play. Furthermore, if it became statistically normal for optic nerves to produce sensations of pleasure that stimulate philosophical creativity, it would not follow from this that producing such sensations is what optic nerves are supposed to do.

If this is right, then SN-accounts of proper function will not do for our purposes. But what of etiological analyses? We should note first that what initially suggests itself as the most intuitive etiological analysis is also most likely to be unacceptable to a naturalist. The analysis I have in mind is this: the (or a) proper function of X is to F just in case X was designed to F. Though this analysis is quite intuitive, it commits naturalists to thinking that natural objects, one and all, lack proper functions (again, assuming that the methods of science alone do not reveal the existence of a cosmic designer). Thus, insofar as the naturalist will want a unified account of proper function for both natural objects and artifacts, and insofar as she will want to say that at least some natural objects have proper functions, she is committed to saying that it is some sort of natural etiology that confers proper function – not an etiology that somehow involves a designer. As it turns out, the etiology of this sort that is universally invoked by naturalistic defenders of proper function is reproductive, or evolutionary etiology. Thus, whereas no one explicitly defends an SN-account of proper function, everyone who explicitly defends a naturalistically acceptable account of proper function defends a broadly evolutionary account. I will discuss only one such account – one proposed and defended independently by Karen Neander and Ruth Millikan.[27] This is not the only account of proper function in the literature, but it is the most prominent and the objections to it carry over to the others.

The full statement of Millikan's account of proper function occupies two chapters in *Language, Thought, and other Biological Categories*. But in her "In Defense of Proper Functions," we find the following summary of her view:

> Putting things very roughly, for an item *A* to have a function *F* as a "proper function", it is necessary (and close to sufficient) that one of these two conditions should hold. (1) *A* originated as a "reproduction" (to give one example, as a copy, or a copy of a copy) of some prior item or items that, *due* in part to possession of the properties reproduced, have actually performed *F* in the past, and *A* exists because (causally historically because) of this or these performances. (2) *A* originated as the product of some prior device that, given its circumstances, had performance of *F* as a proper function and that, under those circumstances, normally causes *F* to be performed by *means* of producing an item like *A*. Items that fall under condition (2) have "derived proper functions", functions derived from the devices that produce them.[28]

Though this statement does not include all of the subtleties of her more precise statement of the view, it will suffice for present purposes. Neander's account, on the

other hand, is much more simply stated. She writes (under the assumption that the unit of selection in natural selection is a genotype):

> It is the proper function of an item X of an organism O to do that which items of X's type did to contribute to the inclusive fitness of O's ancestors, and which caused the genotype, of which X is the phenotypic expression, to be selected by natural selection.[29]

Before turning to objections, it is important to note that both Millikan and Neander deny that their theories are intended to *analyze* the *ordinary* concept of proper function.[30] Millikan eschews analysis altogether, claiming that she is giving a "theoretical definition" of a "technical term" that is of interest because of its usefulness in solving certain problems and not because of any relation it bears to ordinary notions of function or purpose.[31] Neander, on the other hand, embraces analysis but (a) denies that giving an analysis involves giving necessary and sufficient conditions and (b) claims that the concept of proper function that she is analyzing is a technical concept in biology, not our ordinary concept.[32] Despite initial appearances, then, it is doubtful that either Millikan or Neander means to be giving an account of proper function *as it is understood in the present chapter*. Thus, the objections I shall be raising should not at all be taken as objections to the projects that Millikan and Neander intended to engage in. Rather, they should be taken simply as objections to the idea that their accounts of proper function in their sense(s) of the term could serve as adequate accounts of proper function in my sense of the term.

That the Millikan/Neander account will not do as an account of proper function in the ordinary sense can be seen by considering the now familiar example of reproductive clay crystals. In *The Blind Watchmaker*, Richard Dawkins explains how natural selection could operate on a population of clay crystals.[33] He invites us to imagine a kind of clay that improves its own chances of being deposited by damming up streams. The streams form shallow pools which then dry up. The clay dries and is blown away as dust, only to be deposited in other streams. The new crystals reproduce themselves and dam up their respective streams. And the process begins anew. Here, we have all of the elements of natural selection: the crystals reproduce, display random variation in the traits they possess, pass on traits to their descendants, and sometimes display traits that are either adaptive or maladaptive. But, as Mark Bedau also indicates in his own discussion of this example, no one would want to say that there is any metaphysically important sense in which the crystals are *supposed* to dam up streams.[34] The reproductive history of an item makes no difference with respect to what it is supposed to do.

Granted, crystals are not biological organisms, so strictly speaking, though we do have a counter-example to Millikan's account (construed as an account of proper

function in the ordinary sense), we do not have a counter-example to Neander's. Still, the example casts doubt on Neander's account. *If*, as we all want to say, the reproductive history of a population of clay crystals doesn't tell us anything at all about how those crystals are supposed to behave, why should we think that the reproductive history of a population of biological entities tells us anything at all about how those biological entities are supposed to behave? There seems to be no relevant difference between the two cases.

We might put this point another way: Suppose we concede (as seems correct) that at least Neander's view provides a materially sufficient condition for something's having a proper function in the ordinary sense. That is, perhaps it is in fact the case that everything that satisfies her account also has a proper function. Still, that doesn't guarantee that her account is adequate, for it doesn't guarantee that there is any metaphysically or epistemically important connection between the properties by virtue of which something satisfies her account and the properties by virtue of which it counts as having a proper function. Furthermore, there is reason for thinking that in fact there is no important connection between the two sorts of properties. One reason is just the fact that clay crystals can exemplify many of the most important properties (namely, those involved in a natural-selective history) that Neander thinks guarantee proper function without themselves having a proper function. Another reason is that we can imagine cases where proper function and reproductive history might diverge. For example, we might imagine a world in which an incompetent deity creates biological organisms that evolve in ways contrary to the intentions of that deity. In such a case, it seems clear that whatever we are inclined to say about what (if anything) counts as proper functioning for such organisms, we will not say that there is an objective, metaphysically privileged sense in which the organisms are supposed to function in the way that they evolved to function. Perhaps we will say that they are supposed to function in the way they were designed to function, or perhaps we will say that in this case there is no such thing as proper function for the organisms in question. But we will not say that the evolutionarily determined way of functioning is the objectively proper way.

It is trendy now to object to imaginary examples simply by appeal to the popular idea that conceptual analysis is a misguided project. But I find the objection unpersuasive. Even if conceptual analysis is a misguided project, I see no reason to doubt the evidential value of imaginary examples, especially when the examples are not highly contrived. But for those who insist on rejecting imaginary examples, we may also provide real examples where reproductive history and proper function seem to diverge, thus casting doubt on the idea that Neander has even provided a materially sufficient conditions for something's having a proper function. Consider genetically engineered organisms. Will it be the objective, metaphysically privileged, proper function of a structure altered by engineering to display the reproductively established trait, the engineered trait, or neither? If one of the latter two, then there

are cases where proper function and reproductive history *do* diverge and we have immediate counter-examples to Neander's view. If the former – if engineering always produces malfunction – then questions arise about what we should say about cases where an engineered trait becomes reproductively established. It seems clear that in such cases we should not say that a genuine malfunction has become a proper function. But if we should not say that, then, again, there are (or could be) cases where reproductive history and proper function diverge.

It should be clear that the problem here is quite general. Etiology might guarantee a kind of weak normativity. It might guarantee that there is a weak sense in which a thing is supposed to perform a certain function. But etiology alone cannot guarantee any sort of strong, metaphysically privileged sense in which a thing is supposed to perform a certain function. It simply doesn't follow from the fact that a thing has a certain kind of causal history that it is in some metaphysically important way supposed to do whatever it is that its causal history has determined it to do.

If this is right, then we seem forced to the conclusion that the methods of science alone cannot reveal that anything has a proper function. Science can reveal facts about statistical normality; it can also reveal facts about reproductive history. But, as we have seen, there is no obvious epistemic or metaphysical connection between such facts and facts about proper function. Thus, there is no reason to think that in revealing such facts, science reveals facts about proper function. Furthermore, there are no naturalistically acceptable alternatives – which is just to say that there are no *other* scientifically accessible facts that naturalists might reasonably take as indicative of proper function. Hence, science cannot reveal proper function.

So far, I have been assuming that the naturalist will want a unified account of proper function – one that will cover both natural objects and artifacts. But, strictly speaking, I have said nothing to rule out the possibility that (a) there are no biological proper functions but (b) some artifacts have proper functions by virtue of various intentions and goals had by their human designers. If this is right, then since human intentions and goals are accessible to science, some proper functions are accessible to science after all.

But there is a problem here. Intentions and desires are accessible to science only insofar as they are physical states of biological objects (i.e. brains or parts of brains) or perhaps artifacts (such as computers). But, as earlier arguments in this section showed, we need biological proper function to render belief in mind-independent biological objects naturalistically acceptable. What this means, then, is that whatever a naturalist thinks about intentions and desires, she cannot think that they are physical states of persisting things that have their persistence conditions independently of human mental activity. In other words, she must give up the commonsense view that mental activity can only take place in a persisting, mind-independent subject. That by itself is not an objection to the view, but it *is* a reason for thinking that our best science cannot find anything in nature that satisfies the commonsense conception of an

intention or a desire. But it is intentions and desires *as commonly conceived* that are widely seen as being able to confer proper functions. Who knows whether intentions and desires under the revised conception will plausibly be able to do so? Indeed, who knows whether our best science will even leave room for intentions or desires? Thus, at least in the absence of further information from the field of cognitive science, it seems clear that if it is naturalistically unacceptable to believe in biological proper functions (and hence mind-independent biological objects), it will be naturalistically unacceptable to believe in *any* proper functions. And this, of course, entails that belief in mind-independent material objects generally is naturalistically unacceptable.

I conclude, then, that (3.1) is false. But, again, apart from the truth of (3.1) I see no other way to explain how we acquire justified beliefs about classification short of saying that facts about how objects are to be classified depend importantly upon our mental activity. Thus, I conclude that we cannot acquire naturalistically acceptable beliefs about classification.

In this chapter, I have argued that belief in mind-independent material objects depends for its justification upon our ability to acquire the justified belief that modal properties are exemplified in at least some of the regions of space-time that we take to be occupied by material objects, and I have argued that beliefs of this latter sort depend for their justification upon our ability to acquire justified classificatory beliefs. Furthermore, I have shown that the most promising suggestion for how a naturalist might be able to be justified in classificatory judgments is a failure, and I have observed that if this suggestion fails, there is good reason to believe that any other will fail as well. Thus, I conclude that belief in mind-independent material objects is incompatible with commitment to naturalism.

One might be tempted to respond by saying that non-naturalists are in a similar boat. It is very difficult for *anyone* to tell how we acquire justified classificatory beliefs; the problem is not peculiar to naturalism. However, there is a relevant difference between the position of a naturalist and the position of a non-naturalist with respect to questions about the justification of our classificatory beliefs. Non-naturalists are free to say that classificatory beliefs are justified a priori; and they can furthermore observe that the mystery of how such beliefs are justified is no different from the mystery of how *any* beliefs might be justified a priori. Naturalists, on the other hand, are not free to make this move. Naturalists are committed to thinking that classificatory beliefs admit of *a particular kind of justification* – namely, scientific. Thus, showing that there is good reason to doubt that such beliefs could be justified through the methods of science provides reason for naturalists to be skeptical where non-naturalists need not be skeptical.

As I indicated earlier, some naturalists will be content to abandon the ontology of mind-independent material objects. Some philosophers are attracted to stuff ontologies; and many of the same philosophers are attracted to either conventionalism or conceptualism. Other philosophers are happy with a

thoroughgoing agnosticism about ontology (which is, perhaps, the more appropriate stance in light of what I have argued about the status of our knowledge of the modal character of the world). Nevertheless, many have the strong conviction that it is much more reasonable than not to believe that there really are some things (e.g. people and plants) that last over time and that have their persistence conditions independently of human mental activity and semantic conventions. If the arguments in this chapter are sound, however, those who share this conviction ought to reject naturalism.

Notes

1 Work on this project was supported by a University of Delaware General University Research Grant and by a grant from the Pew Evangelical Scholars Program. I also have an embarrassingly long list of people to thank for advice. Fred Adams, Chris Boorse, Jeff Brower, Tom Crisp, Jeff Jordan, Alvin Plantinga, Bill Ramsey, David Silver, Jim Stone, and Dale Tuggy all read early drafts and provided many helpful comments. A version was also read at the Fourth Annual Mighty Midwestern Metaphysical Mayhem Conference at the University of Notre Dame. Conversations with the participants in that conference – particularly David Chalmers, Brie Gertler, John Hawthorne, Larry Lombard, Brian McLaughlin, and Laurie Paul (as well as others already mentioned) – were extremely helpful. Finally, I am especially grateful to Michael Bergmann, Trenton Merricks and Joel Pust, who were subjected to multiple drafts and offered many very useful suggestions. I would also like to thank Marty Strachan, my research assistant during the summer of 1998, for valuable information on the topic of conventionalism.

2 For example: D. M. Armstrong, "Naturalism, Materialism, and First Philosophy," reprinted in *Contemporary Materialism*, ed. Paul Moser and J. D. Trout (London: Routledge, 1995), 35–47.

3 For example: Michael Devitt, *Coming to Our Senses* (Cambridge, Mass.: Cambridge University Press, 1996); W. V. Quine, "Naturalism; Or, Living Within One's Means," *Dialecitca* 49 (1995): 251–61; W. V. Quine, "Epistemology Naturalized," in *Ontological Relativity and Other Essays* (Cambridge, Mass.: Harvard University Press, 1969), 69–90; and Alex Rosenberg, "A Field Guide to Recent Species of Naturalism," *British Journal for the Philosophy of Science* 47 (1996): 1–29.

4 For example: Jerrold Katz, *Realistic Rationalism* (Cambridge, Mass.: MIT Press, 1998); Jean Hampton, *The Authority of Reason* (New York: Cambridge University Press, 1998); and Barry Stroud, "The Charm of Naturalism," *Proceedings and Addresses of the American Philosophical Association* 70 (1996): 43–56.

5 W. V Quine, "On What There Is," in *From a Logical Point of View* (Cambridge, Mass.: Harvard University Press, 1961).

6 Quine, "Naturalism; Or, Living Within One's Means"; Michael Ruse, *Evolutionary Naturalism* (London: Routledge, 1995); Mark Bedau, "Naturalism and Teleology," 23–51 in S. Warner and R. Wagner, eds., *Naturalism: A Critical Appraisal* (Notre Dame: University of Notre Dame Press, 1993).

7 Armstrong, "Naturalism, Materialism, and First Philosophy"; Hilary Kornblith, "Naturalism: Both Metaphysical and Epistemological," 39–52 in P. French, T. Uehling and H. Wettstein

(eds.), *Midwest Studies in Philosophy, Volume 19, Philosophical Naturalism* (Notre Dame: University of Notre Dame Press, 1994).

8 John Dewey, Sidney Hook, and Ernest Nagel, "Are Naturalists Materialists?" *Journal of Philosophy* 42 (1945): 515–30; William Dennes, "The Categories of Naturalism," 270–94 in Y. Krikorian (ed.), *Naturalism and the Human Spirit* (New York: Columbia University Press, 1944).

9 Wilfrid Sellars, "Empiricism and the Philosophy of Mind," in *Science, Perception, and Reality* (London: Routledge & Kegan Paul, 1963), 173.

10 Michael Rea, *World Without Design: The Ontological Consequences of Naturalism* (manuscript in preparation). This view of naturalism is also accepted by Peter Forrest. See Peter Forrest, *God Without the Supernatural* (Ithaca, N.Y: Cornell University Press, 1996), 89.

11 Note that if facts about what changes a thing can and cannot survive depend upon God's beliefs or attitudes, or upon the beliefs and attitudes of some other non-human being, that thing will still count as mind-independent as I am using the term. Hence the imprecision. But the term is convenient, and the problems are negligible in the present context, so I will continue to use the term in the way just described.

12 See, for example, Thomas Blackson, "The Stuff of Conventionalism," *Philosohpical Studies* 68 (1992): 65–81; Michael Jubien, *Ontology, Modality, and the Fallacy of Reference*, (Cambridge: Cambridge University Press, 1993); and Alan Sidelle, *Necessity Essence and Individuation* (Ithaca, N.Y: Cornell Unviersity Press, 1989). Arda Denkel has recently argued that it is impossible to have matter without objects (Arda Denkel, "Matter and Objecthood," *Dialogue* 28 (1989): 3–16). But see also the replies by Laycock and Sidelle (Henry Laycock, "Matter and Objecthood Disentangled," *Dialogue* 28 (1989): 17–21; and Alan Sidelle, "Formed Matter Without Objects: A Reply to Denkel," *Dialogue* 30 (1991): 163–71.)

13 Rudolf Carnap and Allan Gibbard both seem to be conceptualists. See Rudolf Carnap, *Meaning and Necessity* (Chicago, Ill.: University of Chicago Press, 1947) and Allan Gibbard, "Contingent Identity," reprinted in Michael Rea (ed.), *Material Constitution: A Reader* (Lanham, Md.: Rowman & Littlefield, 1997). Alan Sidelle defends in detail a version of conventionalism independently described but not defended by Paul Coppock. See Paul Coppock, Review of *Reference and Essence* by N. Salmon, *Journal of Philosophy* 81 (1984): 261–70 and Sidelle, *Necessity, Essence, and Individuation*.

14 But see William Dembski, "Naturalism and Design" in this volume.

15 By "modal realism" I just mean the denial of modal anti-realism as it is characterized above. Modal realism is not to be understood as the thesis (endorsed by David Lewis and others) that possible worlds are concrete objects spatiotemporally disconnected from our world.

16 Michael Rea, "The Problem of Material Constitution," *Philosophical Review* 104 (1995): 525–52.

17 The example is from David Wiggins, "On Being in the Same Place at the Same Time," reprinted in Michael Rea (ed.), *Material Constitution: A Reader.*

18 Note that the kind of classification under discussion here is metaphysical rather than, say, biological or physical. Few people nowadays think that biological classifications are metaphysically informative. Note furthermore that, on this view, it is possible to be nominally a K without being classifiable as a K, but not vice versa. For a defense of my views about kind-membership, see my "Constitution and Kind Membership," forthcoming in *Philosophical Studies*. Nothing of substance turns on the assumption that there are two ways of belonging to a kind; but it would be inordinately complicated at best to discuss the points that will follow

in completely theory-neutral terms.

19 Peter van Inwagen, *Material Beings* (Ithaca, N.Y.: Cornell University Press, 1990).

20 See, for example, John Dupré, *The Disorder of Things* (Cambridge, Mass.: Harvard University Press, 1993) and W. V Quine, "Whither Physical Objects?" pp. 497–504 in R. S. Cohen, P. K. Feyerabend, and M. W. Wartofsky (eds.), *Essays in Memory of Imre Lakatos* (Dordrecht: D. Reidel, 1976).

21 Crawford Elder defends views that are in some ways suggestive of this one (though not nearly enough to justify attributing the view to him). Elder maintains that the essential properties of a thing are learned by learning facts about its distinctive *nature;* he also maintains that the proper functions of things are essential to them (see Crawford Elder, "An Epistemological Defense of Realism about Necessity," *Philosophical Quarterly* 42 (1992): 317–36, and Crawford Elder, "On the Reality of Medium-Sized Objects," *Philosophical Studies* 83 (1996): 191–211). If he is right, then perhaps one implication of his views is that we acquire justified classificatory beliefs by way of acquiring justified beliefs about proper function. However, I hesitate to attribute this sort of view to Elder since (a) he doesn't explicitly defend it, (b) both of his papers presuppose that we can make classificatory judgments in arguing for their respective conclusions, and (c) the second of the two papers also presupposes that modal intuitions (such as intuitions about the essentiality of origins) can help to establish facts about the essential properties of material objects.

22 I assume that systems are wholes; hence the objects "in" the system are just parts of it. Note that if function ascriptions are indeed system relative, then to say that some xs are functionally organized entails that they are parts of a whole. Thus, we have a quicker route to a conclusion – that being functionally organized is sufficient for being parts of a whole – that I defended on purely intuitive grounds on p. 354 of Michael Rea, "In Defense of Mereological Universalism," *Philosophy and Phenomenological Research* 58 (1998): 347–60.

23 Here I assume that if A-facts supervene on B-facts, and if science can reveal A-facts, then it can also reveal B-facts. I do not assume the converse – that if science can reveal B-facts then it can also reveal A-facts. I also do not assume the general principle that if way of knowing W can reveal A-facts, and if A-facts supervene on B-facts, then way of knowing W can reveal B-facts. The falsity of this general principle can readily be seen by substituting "introspection" for "W", "facts about mental phenomena" for "A-facts," and "facts about brain-chemistry" for "B-facts" (I thank Trenton Merricks for this point).

24 This is in fact a rough summary of Larry Wright's account of what it is for something to have a *function* (see Larry Wright, "Functions," *The Philosophical Review* 82 (1973): 139–68. For counter-examples, see Christopher Boorse, "Wright on Functions," *The Philosophical Review* 85 (1976): 70–86, and Elizabeth Prior, "What is Wrong with Etiological Accounts of Biological Function?" *Pacific Philosophical Quarterly* 66 (1985): 310–28). Nevertheless, many of those following Wright have adopted roughly this sort of view to explain what it is for something to have a *proper function*. In fact, most of those who write on proper functions tend to conflate the notions of function and proper function (see, for example, Ruth Millikan, "In Defense of Proper Functions," *Philosophy of Science* 56 (1989): 288–302, especially p. 293; Karen Neander, "Functions as Selected Effects: The Conceptual Analyst's Defense," *Philosophy of Science* 58 (1991): 168–84, esp. p. 173; and Michael Levin, "Plantinga on Functions and the Theory of Evolution," *Australasian Journal of Philosophy* 75 (1997): 83–98).

25 Alvin Plantinga, *Warrant and Proper Function* (New York: Oxford University Press, 1993), 199. Plantinga also claims that the analysis is suggested, though not explicitly stated, in John

Pollock's "How to Build a Person," in James Tomberlin, (ed.), *Philosophical Perspectives, volume 1: Metaphysics* (Atascadero, Calif.: Ridgeview, 1987), 109–54.

26 Plantinga, *Warrant and Proper Function*, 200–1.

27 Karen Neander, *Abnormal Psychobiology* (Ph.D. diss., La Trobe University, 1983); Neander, "Functions as Selected Effects"; Ruth Millikan, *Language, Thought, and Other Biological Categories* (Cambridge, Mass.: MIT Press, 1984); Millikan, "In Defense of Proper Functions". Among those who extend their view to include artifacts are Paul Griffiths, "Functional Analysis and Proper Functions," *British Journal for the Philosophy of Science* 44 (1993): 409–22; Crawford Elder, "On the Reality of Medium Sized Objects"; and Beth Preston, "Why is a Wing Like a Spoon? A Pluralist Theory of Functions," *Journal of Philosophy* 95 (1998): 215–54.

28 Millikan, "In Defense of Proper Functions," 288–9.

29 Neander, "Functions as Selected Effects," 174.

30 Even noting this, however, does not entirely remove the possibility of being misunderstood. For example, John Post accuses Alvin Plantinga of misreading Millikan; but, ironically, the accusation itself rests on a misreading of Plantinga (see John Post, "Critical Notice of Ruth Millikan's *White Queen Psychology and Other Essays for Alice," Philosophy and Phenomenological Research* 58 (1998): 233–7). Post claims that "Plantinga…badly misreads [Millikan] as attempting an analysis, then tries to counter-example accordingly" (233). In fact, however, it is Post, not Plantinga, who is guilty of poor exegesis. Plantinga never suggests that Millikan is attempting an analysis, and he explicitly acknowledges that "her project is not as it stands directly relevant to the question at issue: the question whether there is available a naturalistic understanding or analysis of proper function." (Plantinga, *Warrant and Proper Function*, 201). He goes on to say that "Millikan's account…is subtle and challenging (and formidably difficult), and it is worth looking to see whether in fact it provides the materials for an adequate or accurate analysis of the notion of proper function" (202). But, of course, in light of his earlier remarks it is abundantly clear that Plantinga has not made the mistake of thinking that Millikan was actually attempting to provide "the materials for an adequate or accurate analysis of the notion of proper function."

31 Millikan, *Language, Thought, and Other Biological Categories*, 17; "In Defense of Proper Functions," 289–93.

32 Neander, "Functions as Selected Effects," especially pp. 168–73.

33 Richard Dawkins, *The Blind Watchmaker* (New York: W. W. Norton & Company, 1996), 150–4.

34 Bedau, "Naturalism and Teleology," 36–42.

6 Naturalism and the mind

Charles Taliaferro

J. L. Mackie developed an argument against moral realism he called "the argument from queerness."[1] He contended that, given a materialist understanding of the cosmos, the existence of objective values would be queer or, to use less colorful language, objective values would be unique and unlikely. "If there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe."[2] Mackie did not conclude that the strange and unlikely has happened, but instead that there are no objective values. The argument has such currency that it is even an entry in the *Penguin Dictionary of Philosophy* under the heading "queerness." The argument runs along the same lines as Mackie's reasoning in *The Miracle of Theism*. Following Hume, Mackie argues that theism is so unlikely it seems like a miracle that anyone believes it; the existence of God would be "queer" given everything else we know about the cosmos, and thus it is best denied. A yet further, similar argument has been advanced by many philosophers about consciousness.

Some philosophers contend that the existence of consciousness is itself an entity, quality or relation of a very strange sort, a reality that seems utterly different from anything else in the universe. While the universe is fundamentally physical in its origin and constitution, consciousness seems so radically different that we are disposed to think of it as something deeply mysterious, perhaps even something non-physical. Given a general physicalist outlook, contemporary philosophers are reluctant to posit consciousness as some elusive, immaterial object. An overriding physicalist framework leads philosophers either to eliminate consciousness and the mental, lest these entities become untamed emergent items, or it leads philosophers to acknowledge consciousness and related mental states, but to treat these as physical. Paul Churchland describes the prevailing physicalist orientation of much philosophy of mind:

> Most scientists and philosophers would cite the presumed fact that humans have their origins in 4.5 billion years of purely chemical and biological evolution as a weighty consideration in favor of expecting mental phenomena to be nothing

> but a particularly exquisite articulation of the basic properties of matter and energy.[3]

This observation fits with Jaegwon Kim's estimation of the current field. Kim points out that the goal for many philosophers is to accommodate physicalist, background assumptions about the nature of the world while not explaining away the mental altogether and eclipsing vital, human values:

> The shared project of the majority of those who have worked on the mind–body problem over the past few decades has been to find a way of accommodating the mental within a principled physicalist scheme, while at the same time preserving it as something distinctive – that is, without losing what we value, or find special, in our nature as creatures with minds.[4]

The goal is either to accommodate consciousness, minds, values *et al.* as denizens of a physical world or to eliminate them. Either way, materialism is the order of the day. As William Lycan writes in *Consciousness and Experience,* "Few theorists question the eventual truth of materialism."[5]

Naturalism admits of a range of definitions and scales. There are, for example, strong versus weak naturalists, radical and agnostic naturalists, and so on. Sometimes naturalism is simply another name for a thoroughgoing form of physicalism, while at other times naturalism is simply *any* view of the world that is incompatible with supernaturalism. In a volume dedicated to raising difficulties for naturalism, I shall point to the problem of incorporating the mental in a physicalist schema, and I shall then endeavor to clear the ground for a view of human nature that is resisted by all physicalists and virtually all naturalists. If this non-materialist, non-naturalist alternative may resolve the problems facing materialism and naturalism, and if this alternative may be seen to have some overriding plausibility, then this essay will have done its part in the overall case against naturalism. The non-materialist, non-naturalist philosophy of mind I shall defend is a form of mind–body dualism. Daniel Dennett's comment on dualism captures the prevalent, naturalist disdain toward dualism. In *Kinds of Minds* he describes dualism as the "view that minds are composed of some nonphysical and ultimately mysterious stuff . . . relegated to the trash heap of history, along with alchemy and astrology."[6] My case for dualism will not lean on alchemy and astrology, but it will be in concert with philosophical theism, as sketched in a final section of this paper.

Not long ago, philosophers did not believe that a person is "a particularly exquisite articulation of the basic properties of matter and energy." It was commonplace, rather, to believe that we live in a theistic cosmos and that human beings are an exquisite articulation of the Creator of the cosmos. A range of philosophers today link the case

for dualism and theism.[7] I shall take this link seriously. Essentially, my goal is to develop a position similar to George Mavrodes' in his reply to Mackie's argument from queerness. In "Religion and the Queerness of Morality"[8] Mavrodes argues that objective values look less "queer" if seen against a theistic back-drop than they do if seen in light of Mackie's naturalism. Following Mavrodes' lead, I contend that the case for dualism is far more promising once debate in philosophy of mind is carried out in a broader metaphysical context.

Problems with eliminative materialism

Eliminative materialism is not as popular as it was some decades ago. A major problem has been the task of developing a version of eliminative materialism that is not self-refuting or self-contradictory. Some eliminativists appear to be in the unenviable position of claiming to believe that there are no beliefs. Another difficulty is the problem of being able to accommodate human reasoning. A further worry still is that eliminativism is flatly refuted by experience.[9]

The next section of this paper will consider the problem of treating consciousness and experience within a form of reductive or identity materialism. Insofar as experience does have the characteristics I shall argue that it does, this counts as a problem for eliminativism as well. But here I wish to consider a problem with the way eliminativism is articulated in the literature. Eliminativism has been advanced in a fashion that appears to leave little unsettled in our everyday discourse and our understanding of social relations and the like. I believe this is deeply problematic. Consider, for example, Paul Churchland's *The Engine of Reason, the Seat of the Soul*.

At the outset, *The Engine of Reason, the Seat of the Soul* appears non-eliminative. Churchland asks: "How *does* the brain work? How does it sustain a thinking, feeling, dreaming self?"[10] This seems to involve a straightforward acknowledgment of thinking, feeling, dreaming, and of a unitary subject that engages in each. Indeed, some of Churchland's depiction of the findings of the brain sciences seem compatible with a philosophy of mind and body that is either a form of *non*-eliminative materialism or even dualism:

> [W]e are now in a position to explain how our vivid sensory experience arises in the sensory cortex of our brains: how the smell of baking bread, the sound of an oboe, the taste of a peach, and the color of a sunrise are all embodied in a vast chorus of neural activity.[11]

Reference to how a psychological phenomena "arises" suggests a distinction between the psychological and mental, and talk of "embodiment" seems as natural for noneliminative materialists as it is for dualists who uphold the metaphysical

distinction between the mental and physical and yet insist upon their causal interaction. But this open-ended description of the mind and body is soon put aside.

Churchland has elsewhere construed dualism as "mankind's official view of itself" and so it is often Churchland's foil in the course of his arguments for a thorough materialism.[12] But Churchland challenges not only dualism, but non-reductive and identity materialism as well. One of the themes of Churchland's book is that our customary theory or working assumption about ourselves needs to be radically revised in light of the natural sciences. Our "official" and unofficial, ordinary conceptions of ourselves is mistaken:

> Is our basic conception of human cognition and agency yet another myth, moderately useful in the past perhaps, yet false at edge or core? Will a proper theory of brain function present a significantly different or incompatible portrait of human nature? . . . I am inclined toward positive answers to all of these questions.[13]

The challenge to folk psychology comes in stages. At certain points, a unitary subject distinct from the brain is the target and the properties that would customarily be attributed to this subject are attributed to parts of the brain. Churchland comments on vision:

> But who . . . can be watching this pixilated show? The answer is straightforward: no one. There is no distinct "self" in there, beyond the brain as a whole. On the other hand, almost every part of the brain is being "watched" by some other part of the brain, often by several other parts at once.[14]

But then the claim that we folk customarily watch things – when this is interpreted in terms of sensations, concepts, and beliefs – is itself critiqued.

Churchland assumes that the average reader, or newcomer to the field, is not an eliminativist when it comes to sensations, beliefs, perceptions, desires and preferences:

> You came to this book assuming that the basic units of human cognition are states such as thoughts, beliefs, perceptions, desires, and preferences. That assumption is natural enough: it is built into the vocabulary of every natural language. . . . These assumptions are central elements in our standard conception of human cognitive activity, a conception often called "folk psychology" to acknowledge it as the common property of folks generally. Their universality notwithstanding, these bedrock assumptions are probably mistaken.[15]

This is a radical claim; it takes us some distance beyond the customary attack on dualism. Churchland claims that the disposal of such folk "bedrock assumptions" and the building up of an alternative, scientifically-informed understanding of human nature will lead us to new insights into moral life. New scientific discoveries can "set us free, and allow us to achieve a still higher level of moral insight and mutual care."[16] He endorses a form of moral realism and indicates his preference for an Aristotelian as opposed to Kantian ethic. Brain and physiological testing provide Churchland with evidence that moral judgments on specific cases comes prior to our grasping and applying moral rules.

Churchland proposes that a successful, eliminative account of persons will not leave us with some "cold" or soulless view of ourselves:

> One's first impulse, perhaps, is to see the vocabulary and framework of a general theory of the brain as something alien and cold. But it will not be alien if it depicts all of us, at last, as we truly are… Whatever the distractions, we must continue to exercise our reason. And whatever the temptations we must continue to nurture our souls. That is why understanding the brain is so supremely important. It is the engine of reason. It is the seat of the soul.[17]

Churchland does not see that eliminativism will threaten "the soul" and, perhaps to put on display his settled view on the compatibility of eliminativism and his philosophy of personal relations, he includes several personal allusions in *The Engine of Reason, the Seat of the Soul*. There is, for example, a photograph of his daughter and her "soul-mate," which is then analyzed in terms of retina-visual cortex interaction.[18] And there is figure 7.1, an MRI image of Patricia Churchland's brain.[19] Paul Churchland comments: "This particular brain is in fact well known to me via more conventional informational pathways. It is the brain of my wife and colleague, Patricia Churchland, and it is very dear to me."[20]

But regrettably, it is not very clear how Churchland can succeed both in advancing eliminativism and redeeming the promise that it will assist us in an effort to "nurture our souls" and care for one another. Obviously the brain sciences can assist us insofar as they enable us to treat physiological illness and the organic basis for psychological dysfunction. They can provide the non-eliminative philosopher with insights as to the material underpinnings of our mental life. But it is not clear how we can make use of notions life "assist," "care" and so forth, if we must shed bedrock assumptions about thoughts, beliefs, perceptions, desires and preferences. In most, if not all cases, "care," "compassion," and "love" appear to make sense only within a conceptual context of beliefs and desires. At a minimum, love between persons seems to involve beliefs about oneself, beliefs about the beloved, and some desire for one another's well-being. In *The Engine of Reason, the Seat of the Soul* we are not given a clear guide as to how what, in

folk language, we refer to as "love," "care," "compassion," and so on, can be given any purchase in an eliminative world.

Notwithstanding the pride of place that Churchland gives to the moral benefits of the position he defends, I think the overall, eliminative project gives rise to similar worries about whether the brain (or "enbrained body," to use an awkward phrase) as Churchland describes it can indeed function as an engine for reason. In my view, it is difficult to conceive of how any reasoning can survive in an eliminativist framework. In a world that eliminates the intentional world of folk psychology, how can one make sense of reasoning? In a standard case of reasoning, one accepts a conclusion in virtue of believing the premises of an argument and acting on one's grasp of the nature of entailment. Jane Doe accepts conclusion B, because she believes that 'If A then B' and she believes there is compelling evidence that A. I am skeptical about whether reasoning can be given any intelligible place in our view of what we do when we undertake philosophical reflection (or any other norm-governed reflection in which conclusions are derived on the basis of beliefs) if one does away with beliefs – or, if one acknowledges the existence of beliefs and yet gives them no role to play in a causal explanation of why, say, Jane Doe reasons as she does. Thus, even when Churchland permits (if only as by way of marking a starting point for further analysis) attributions of "beliefs" in describing our activities, I contend that these beliefs are not given a sufficient role to describe and explain (what at least appears to be) our reflection – our coming to hold certain beliefs on the basis of others. Here, I think Churchland's privileging of physics and chemistry causes problems. He is careful to delimit causal accounts to physics and chemistry:

> I do not mean to suggest that social properties are anything more, ultimately, than just intricate aspects of the purely physical world. Nor do I wish to suggest that they are independent causal properties over and above what is captured by physics and chemistry.[21]

But unless physics and chemistry include the stuff of reason and beliefs (which, at least in conventional versions of physics and chemistry, they do not), I do not see how Churchland can do justice to our capacities to reason and argue rationally.[22]

I cited Churchland's claim earlier that a general theory of the brain could not be "alien and cold" so long as "it depicts all of us, at last, as we truly are." But the question that needs to be raised is the prior one of whether eliminativism *can* depict us. For what account may we have of personal identity without taking seriously our beliefs, desires, hopes, loves, and the like? If Churchland is right that a true, general theory of the brain cannot be alien and cold, and the theory of the brain that is at the heart of Churchland's eliminativism seems to leave us in alien and cold waters, then we have reason to think eliminativism is mistaken.[23]

Assuming there is some reason to recognize the existence of beliefs, desires, and other "folk" items, what are the prospects of identifying them as no more than physical processes and states?

Problems with identity materialism

Condensing the problem with materialism into the simplest terms, the chief obstacle to identifying a person's mental life with their body or bodily parts and processes, is that they do not appear to be identical. Identifying the two seems to cut directly against what we (at least seem to) know of our experience and our bodies. On the face of it, conceiving of some mental state like *being in pain* or *being puzzled philosophically* is radically different from conceiving of being in whatever physical states that are posited or described by materialists. It appears that one may conceive of the mental without the physical, and that conceiving of one does not involve or entail the other. The problem is not that the mental and physical appear to be different given our commonplace theory about human nature, but that their apparent difference constitutes a reason for adopting a theory of human nature that does not collapse the two into one thing or kind of thing. As Colin McGinn writes "Consciousness certainly *seems* quite different from a mere brain process. My hearing of a loud bang, say, presents itself as a different *kind* of thing from electrical activity in a certain part of my brain. And thinking about a trip to the beach does not *feel* like the spiking of innumerable neurons in my cortex."[24]

The rationale for differentiating the mental and physical relies upon a widespread method for differentiating all cases of objects and properties. At the very foothills of cognitive development one learns to differentiate objects and states of ourselves by noting how objects may be conceptually identified independently of one another. One does not differentiate oneself from one's care-giver, for example, simply by undergoing an array of sensory experiences, but by conceptually grasping that conceiving of him/her is not the same as conceiving of oneself. It is this conceptual power which, when wedded to visual and other sensory experiences, allows us to perceive the relevant differences.[25] And, a bit later on in our cognitive life, it is because I can conceive of the property of *Being a philosophy professor* without thereby conceiving of the property of *Being the President of the United States* that I recognize these as distinct properties. If someday I discover that there is someone who is both a philosophy professor and the President, I do not thereby discover that the properties are identical but I discover instead that the same person is lucky enough to have both properties.[26]

A related, but I think unsatisfactory, way of putting the above point about differentiating properties (or things in general) is that properties are differentiated when they are describable in terms that are not synonymous. It is because "Being a philosophy professor" and "Being a President of the United States" mean different things that we differentiate professors and presidents. But I believe the descriptions

mean different things because of our power to conceive of the one property without the other; we do not conceive of them differently because of the difference in the meanings of the terms describing them. Besides, the power to differentiate properties and things based on our conceptual powers may operate without descriptive terms. I believe that X (some thing to my right that I have no term for) is different from Fred (where imagine "Fred" names someone but the proper name "Fred" has no descriptive content) because I can conceive of the one without the other.

The appeal to our conceptual powers in differentiating properties is not advanced here as a power that operates independent of our other powers to sense, perceive, reflect on, and remember things. Our ability to differentiate things works in concert with these other powers and is integral to them. And it is on the basis of extensive conceptual investigation that one discovers the relation between various properties, whether these are discoverable a priori (as I believe that some are) or a posteriori, as discovered in experience.

Dualists have advanced a host of thought experiments to capture the differences between the mental and physical. Examine someone's brain states, their body as a whole and any of their parts, while the person is in certain mental states (e.g. he is in pain and also puzzled about philosophy). Does one thereby see any mental states? It appears that one does not. Examining pain as a feeling seems to be one thing and examining the relevant physical states seem to be two different things. The most famous, Cartesian strategy is to argue that the mental can be conceived of without conceiving of anything spatially extended and thus material, and vice versa. A popular, contemporary way of differentiating the mental and physical is by imagining that someone knows all that can be known about the physical world, but that the person lacks the ability to see color. Arguably, if she were to gain visual experience of color she would then grasp something true about the world (e.g. she would know what the color red looks like) that she would not know before.[27] Arguably, this exhibits an important difference between the physical basis or cause of color and other sensory, mental states and those mental states themselves.

In opposing a form of identity materialism, I will not offer a refined analysis of the mental, systematically delimiting the extent and interrelationship of different mental properties and states. I will refer to the mental as a general realm that includes sensing, perceiving, and having intentional attitudes (hoping, loving, hating, remembering, *et al.*). Some philosophers elect to refer to the mental in terms of consciousness (McGinn), experience (Strawson), "something it is like" (Nagel), phenomenally conscious states (Block), and *qualia* (Stubenberg). My principle goal is simply to make trouble for naturalism, and point the way to what I think is a more constructive account of human nature, not to carry out a substantial investigation of the mental and establish a uniform, orthodox system of categories. It will suffice to point out how difficult it is to accommodate some mental properties and states, and leave to another occasion the project of refining our treatment of the mental (e.g. addressing the

question of whether one should posit *intending* as a basic mental category or whether intending can be analyzed in terms of beliefs and desires). I shall adopt what Ned Block in a similar context has referred to as "a liberal terminological policy".[28]

It may will be useful to consider the dualist appeal to experience in the context of weighing four materialist objections. I take these objections from work by Georges Rey, William Lycan and Paul Churchland. While an eliminativist, Churchland has also offered powerful anti-dualist arguments that would clear the way for identity materialism. Like the philosopher John Foster who is an idealist as well as a defender dualism as his next-best philosophy, Churchland defends the identity form of materialism even if his final preference is for eliminativism.

Referential opacity

Some materialists claim that the dualist appeal to experience runs aground because of its failure to recognize referential opacity. The same object may be referred to in different ways and in terms that are not synonymous, and yet there is only the one object being referred to. Consider Georges Rey's analysis of a dualist argument. Rey considers the proposal that the mind or the mental seems to be so radically distinct from the physical that it has the property of being possibly unextended, that is, it might exist without thereby having spatial extension. But my body and bodily states do not seem to have this property. It seems absurd to think one's body (or any physical thing whatever) might have the property of existing without spatial extension. He casts the argument as follows:

(1) My mind has the property of being conceivably unextended.
(2) My body does not have the property of being conceivably unextended.
Therefore (3) My body is not my mind.

Rey analyses this argument as follows:

> Now this argument is patently invalid. The same form of argument could defeat almost any identity claim, consider, for example, the identity of Mt. Everest with the highest mountain on earth. By parallel reasoning, we could argue:
> (1*) The tallest mountain has the property of being conceivably different from Mt. Everest.
> (2*) Mt. Everest does not have the property of being conceivably different from Mt. Everest.
> Therefore: (3*) The tallest mountain is not Mt. Everest.
> . . . By the same reasoning, the morning star couldn't be identical with the evening star, Mark Twain wouldn't be identical with Sam Clemens, and the square of 14 might not be 196! Clearly something has gone wrong.[29]

But his purported analogous arguments do not run to parallel the original dualist one.

If my mind has a property that my body lacks, then the mind is not the same thing as the body. Following the principle of the indiscernibility of identicals (if a=b, then any property of a is a property of b and vice versa) if Mark Twain does have some property *different* from Sam Clemens, then we have two people. But in fact we do not have two people and moreover if the terms "Mark Twain" and "Sam Clemens" *do* pick out the same individual (call this person X) then whatever is true of Mark Twain is true of Sam Clemens. What seems to be in confusion here is a different point, namely *Being called Mark Twain* is a different property from *Being called Sam Clemens*. X has both properties. Similarly the property of *Being called the Evening Star* is different from *Being called the Morning Star*, and *Being the tallest Mountain* is a distinct property from *Being called Mt. Everest* despite the fact that a planet and a mountain each have one pair of properties. *Being the square of 14* is also a distinct property from *Being 196* and the infinitely many other properties had by that number, e.g. *Being the successor of 195, being an even number*, and so on. The appeal to referential opacity does not belie the appearance that the mental and physical are distinct. If one has reason to believe that conceiving of one's mental state does not require conceiving of the physical state it is supposedly identical with, then one has reason to think the two states are indeed two.[30]

Bridging the gap with more science

Some materialists claim that greater scientific awareness will help us see the mental in (or even as) the physical. Consider Churchland's response to a dualist thought experiment referred to earlier in which one imagines looking at the brain but failing to see or thereby identify the relevant mental states. Churchland offers this reply to Leibniz' dualist scenario in which one may (*contra* Leibniz) recognize some mental state such as the taste of peaches in the appropriate physical configuration:

> It remains possible, even granting Leibniz' story, that the taste sensation of a peach is identical with a four-element activation vector in the gustatory pathways. And it remains possible that, should you and I happen to know what vectors constitute what sensations, and should we happen to know where and how to look for those activation vectors, then we might recognize those sensations, from our mitelike perspective, as they go by.[31]

Churchland claims that this spatialization of taste helps shed light on our ability to compare tastes. The reason why tasting pears and apricots is similar is because of the proximity and resemblance between the relevant activation pattern across the four types of tongue receptors as re-represented downstream in one's taste cortex. The peach pattern differs from the apricot pattern by only a few percentage points in each of the four dimensions.[32]

But in *knowing what peaches taste like* and in *knowing the location and character of a four-element activation vector*, are we attending to two properties or activities or just one? I believe that a natural (and plausible) view is that we do not know, nor can we observe, the *constitution* of one by the other, but, rather, we can take note of the link, the interaction or embodiment of the one by and in the other. We do not observe the taste *in* the pattern, nor do we observe the brain pattern *in* reflecting on the taste, but, so it appears, we may well come to conclude that the two are intimately, causally connected such that the differences in taste are causally explained (in part) by the configuration and activity of brain activity. One can conceive of subjective properties like *being such that one tastes peaches* without conceiving of *being gustatory pathway XYZ* (however this is described in the language of current physics and chemistry) and this is an important ground for concluding that the properties are distinct. If this criterion is plausible, then one has grounds for regarding the relevant mental and non-mental, physical properties as distinct. For scientific and other practical purposes, one may well elect to treat the two as a single thing, but it remains a philosophically substantial issue as to whether one should collapse the two metaphysically. This is the point that many dualists have made: correlation is not equivalent to identity (e.g. the older generation of twentieth-century dualists: H.D. Lewis, A.C. Ewing, C.A. Campbell, C.J. Ducasse, *et al*, and "the next generation," e.g. Richard Swinburne, W.D. Hart, Stewart Goetz, *et al.*). Correlation is a necessary but not sufficient condition for identity. Even correlation in all possible worlds does not amount to identity (presumably the properties *being the successor of 5* and *being the smallest perfect number* are distinct but necessarily co-extensive).

Confusing metaphysics and epistemology

Some materialists charge that the dualist appeal to experience fails to take seriously the difference between metaphysics and epistemology. Churchland diagnoses the appeal of dualism as "a conflation between different *ways of knowing* on the one hand, and different *things known* on the other."[33] To bring to light his point, Churchland notes how the same physical state, the location of one's limbs for example, can be known in different ways:

> Here, however, the object of knowledge is exactly the same from both perspectives, the subjective and the objective, and it is something paradigmatically physical: the configuration of your body and limbs.[34]

Other analogies involve access to one's bladder, bowls, stomach, micro-muscles in one's skin, lungs, flushed skin, and so on.

> Such examples can be tripled, quadrupled, and more, but these eight will serve to make the point. The existence of a proprietary, first-person epistemological access to some phenomenon does not mean that the accessed phenomenon is nonphysical in nature. It means only that someone possesses an information-carrying causal connection to that phenomenon, a connection that others lack.[35]

Has Churchland thereby defeated the dualist appeal to subjectivity? I do not think so.

The different ways of knowing the object(s) Churchland cites – by feeling it oneself rather than not feeling it and knowing it in some other way, e.g. observing it scientifically – constitute different things that are knowable. To know the feel of one's flushed cheeks is different – as a feeling – from knowing what it looks like from a few yards away. The point of dualist arguments is that the perspectives of the mental and physical are different sorts of things (where a "thing" is understood broadly to cover properties, activities, processes, and not limited to concrete individual substances). Insofar as Churchland allows that there truly is a difference between the subjective appraisal of one's body, desires, tastes, and knowing about these in other ways, by external observation, say, then he is paying homage to the kind of subjectivity that dualists seek to bring to light. Churchland's reference to "information carrying processes" to cover both first-person observation and the kind that is captured in the language of physics and chemistry appears to override the radical differences in the mental as gleaned, on the one side, in our first-person feeling and subjective appraisal of our own states and attitudes, and, on the other, that which is depicted in physics and chemistry. The distinct reality that is manifested in subjectivity is not exhausted by citing its information content; its content is non-propositional. Even when one is thinking propositionally, I suggest there is a felt awareness of the process (there is *qualia* or what its like to thinking). Apprehending information involves subjective awareness.[36]

William Lycan advances a similar attempt to defuse the case for dualism by claiming that dualists fall prey to an epistemic confusion. He thinks dualists commit what he calls the "stereoptic fallacy":

> The Steroptic [or steroscopic] fallacy of supposing that, because to have a vivid perceptual experience of oneself is nothing like observing the brain of someone else who is having it, having such an experience must be entirely different in nature from any goings-on in the brain that underlie one's having it.[37]

And Lycan's contempt for this is clear: "Even to call this inference a 'fallacy' is to flatter it." He goes on to reprimand those who think otherwise:

> No materialist theory of the mind ever entailed that watching the gray cheesy brain of someone who is having an intense cyan (or whatever) sensation is qualitatively or in any other way like having that sensation oneself. Watching the brain produces gray, cheesy visual sensations in the watcher, subserved by whatever neurophysiology underlies gray, cheesy visual sensations, and gray, cheesy visual sensations are, of course, neurophysiologically and functionally quite unlike the intense deep cyan sensations being held by the subject whose neurophysiology is being watched.[38]

But the problem for materialism is not overlooking the fact that looking at a gray brain will mean looking at something gray. The problem is that if a person claims to have sensations, intense deep cyan sensations, and these are held to be physical, we should be able to locate them as physical states or activities. But which physical property is the very same thing as the intense sensation? If there is the feeling of *being irritated by a fallacy*, which property in the physical sciences is this? Lycan describes materialism, his view, as follows: "If materialism is true, then human beings are large collections of small physical objects and nothing more, ontologically. It follows that any human being could be described, and described completely, in purely scientific terms."[39] But how might one find the scientific term for *being irritated by a fallacy* and other psychological properties?

Arguments from analogy

I note finally two attempts to dismiss the dualist argumentation on the basis that the dualist position is no better than certain inadequate analogies.

Churchland objects that dualists who resist identifying the mental and physical are no different from a stubborn vitalist who assumes that life must be different from what one observes in the natural sciences. Notice first Churchland's depiction of a vitalist claim:

> However closely you might watch these molecular structures folding, unfolding, hooking together, unhooking, and drifting aimlessly around in the soup, it is obvious that you would never observe the impulse of life that urges its growth; you would never observe the telos of life that knows and guides its species-specific development.[40]

Qualia are as useless in providing reasons for dualism, just as this predicament is without any force in favor of vitalism.

The problem is that if vitalism is understood as asserting that there is such a thing as knowing, felt urges, guidance on the basis of some grasp of (or maybe belief in)

purpose and so on, then the difficulty of identifying these with physical processes observed in the natural sciences is very much in place. It is only because the concept of "life" as it is employed by most contemporary biologists seems to be a straightforward non-mental category (defined in terms of cell replication and so on) that there is no philosophical puzzle about the claim to observe life or identify it with physical processes. The question remains, however, whether the dualist position about our own (and some non-human animal) mental lives should be jettisoned as quickly as we have given up thinking of "life" itself in quasi-mental terms. It is one thing to give up the notion that "life" should be understood in mental categories and another to give up the notion that our thinking, believing, feeling, sensing and the like should be understood as *bona fide* mental categories.

Churchland's disregard of qualia comes up in another analogy in which he charges that dualists commit a confusion between types of awareness as one finds in a commonplace manual. Churchland complains about the confused metaphysics in a microwave instruction book. The manual includes this claim:

> The microwaves agitate and vibrate the moisture molecules at such a great rate that friction is created; the friction, in turn, creates heat and the heat causes the food to cook.[41]

Churchland comments:

> This raises a problem: how to connect heat with the rest of what is going on. Here the authors fall back on their prescientific folk understanding of one of the many things that can cause heat: friction! The result is massively misleading to the innocent reader, who is left with the impression that rubbing two molecules together causes heat in the same way that rubbing your two hands together causes heat.[42]

There is a confusion in the manual, but I take issue with Churchland's diagnosis.

The puzzle created by the manual results because it can be read as switching from talk of "heat" as molecules in motion to talk of "heat" as a feeling of warmth, and then the manual appears to revert back talk about molecules in motion. In outlining how a microwave cooks (though not how microwaved food tastes) there is no more need to invoke feelings of warmth than there is a need to invoke vitalistic talk of "urges" when describing the growth of a tree. Instead of playing into Churchland's hands, I believe his pointing to an awkwardness in the manual brings to the fore an awkwardness of entirely dispensing with folk, subjective notions of warmth. I suggest we do wind up with a truncated and unsatisfactory theory of what it feels like to be warm if we were to construct a manual on rubbing hands together using Churchland's data that only

referred to molecular friction. In a sense, then, Churchland's use of the example can be turned around. He seems to acknowledge there is a difference between "heat" when this is engendered by rubbing hands and "heat" when this designates molecules in motion when no sentient being is in the vicinity. Advocates of the dualist appeal to experience would pay him the complement of taking this distinction more seriously, though they would carry it to a conclusion Churchland would reject.

Nonreductive materialism and the limits of the natural world

If one grants a distinction between the mental and the physical as the physical is conceived of by Lycan, Churchland, Rey, Dennett *et al.*, why not simply extend the concept of the physical to cover this stray data? A range of philosophers, including Noam Chomsky, John Searle, and Galen Strawson, seem to challenge the limits set on the physical by standard forms of materialism. Churchland is bent on not allowing entities to escape his materialist net. On proprioception he writes: "There is nothing supraphysical, nothing beyond the bounds of physical science here."[43] And later: "Does the undoubted existence of this unique way of your knowing about your own internal states mean that there is something nonphysical about those states, something that must transcend representation within physical science?"[44] But why not allow that there is something beyond a limited notion of what is physical?

One needs some notion of what *might* be beyond the natural or physical world in order to see what might be at stake in loosening or tightening our concept of what counts as physical. Leopold Stubenberg makes the excellent point that if there is no alternative to materialism, we might just as well treat qualia, subjectivity and experience, as something that will eventually be incorporated into materialism even though we currently have little idea on how this might be accomplished:

> Materialistic science stands unrivaled. The belief that consciousness will force this giant onto its knees may even seem slightly mad.[45]

Colin McGinn also holds that a materialist naturalism is the only real alternative.

> It is either eliminativism or miracles or hidden structure. Absolute noumenalism is preferable to denying the undeniable or wallowing in the supernatural.[46]

McGinn argues for a form of naturalism that posits a hidden structure linking the mental and physical.

I think this approach is sensible. In the absence of an alternative to materialism we should stick it out with materialism, opting either for McGinn's hidden structures or

working to display these structures in the clear light of day. But in staying with materialism we will be resisting not just some of the points made above against identity materialism, but also resisting our (at least apparent) ability to see that the mental and physical are only contingently related. Thought experiments abound in the philosophical literature that cause difficulties for identity and non-reductive materialists who believe that the mental and physical are inextricably bound together (or, using popular, current terminology, that the mental supervenes on the physical). Plausible thought experiments describe cases in which a person has inverted senses (seeing, smelling or tasting in ways that are systematically different from others notwithstanding identical physical conditions), absent qualia (a person lacks a qualia that should be present, given current materialist claims), zombies or superspartans (in which consciousness is missing when, if materialism is right, consciousness should be present), nations instantiate the materialist analysis of mental states and yet do not have the requisite mental states, persons switch bodies without any bodily parts switching, or persons become altogether disembodied.[47] None of these ostensible possibilities seem to be *bona fide* possibilities given most forms of materialism.

One materialist strategy at domesticating these thought experiments is to propose that they represent pictures of worlds in which dualism is true but, as it happens, in this world, dualism is false and thus these are not real possibilities. But this seems to avoid getting at the deeply rooted sense of contingency captured in the thought experiments. Just as my imagining I could be in Rome seems to be grasping a genuine possibility about me in my present circumstances and not just grasping something about me in some different universe, the above thought experiments can be structured so that they reveal what appear to be genuinely contingent relations in this world.[48] Theism seems to provide a framework in which we may understand the relevant contingency and incorporate this contingency in a broader, explanatory framework.[49]

According to theism, the origin and continued existence of the cosmos is in virtue of the intentional activity of an omnipotent, omniscient, good being, God. Theists do not see intentionality as an emergent entity from an utterly mindless background. Behind the emergence of animal and human consciousness is the antecedent consciousness or mind of God. Consciousness is not, then, something queer or anomalous, a 'parlor trick'. So, McGinn puts the alternatives as follows:

> One wants to insist, consciousness cannot *really* be miraculous, some kind of divine parlor trick. It must fit into the natural order of things somehow. Its relation to matter must be intelligible, principled, law-governed. Naturalism about consciousness is not merely an option. It is a condition of understanding. It is a condition of existing.[50]

But in a theistic view of consciousness, there is no parlor trick or discrete miraculous act of God behind the emergence of consciousness. Consciousness emerges from the physical cosmos through an abiding comprehensive will of God that there be a world of physical and non-physical objects, properties, and relations. The relation between matter, energy, consciousness, the laws of space-time, *tout court*, all stem from an overarching, divine activity.

In a theistic metaphysics, the materialist reductive proposal of Dennett and others are reversed. As mentioned at the outset, contemporary philosophy of mind is almost entirely materialist. For many philosophers, the key philosophical project is therefore to locate the mental in the physical world, and this means explaining the mental in terms that are not mental. So, Daniel Dennett privileges explanations of nature that are, at base, non-mental and make no use of mental terms like intelligence and purpose. In his view, to achieve an ultimately satisfactory explanation of human intelligence requires that one not invoke ever increasingly intelligent forces in the universe, but, rather, one accounts for intelligence in terms that ultimately can account for it without invoking intelligence. Dennett contends that accounts that still leave intelligence unexplained are question begging:

> The account of intelligence required of psychology must not of course be question-begging. It must not explain intelligence in terms of intelligence, for instance by assigning responsibility for the existence of intelligence in creatures to the munificence of an intelligent Creator . . .[51]

Dennett's position will thereby feed into the project of securing the unity of science. By explaining intelligence in categories that do not involve it, one can aim to unify the kinds of accounts used in science. Dualists by comparison seem left with intelligence plus all the rest; the mental appears to dangle. "Only a theory that explained conscious events in terms of unconscious events could explain consciousness at all."[52] Georges Rey adopts a similar strategy:

> Any ultimate explanation of mental phenomena will have to be in *non*-mental terms, else it won't be an *explanation* of it. There might be explanations of some mental phenomena in terms of others – perhaps *hope* in terms of *belief* and *desire* – but if we are to provide an explanation of all mental phenomena, we would in turn have to explain such mentaliastic explainers until finally we reached entirely non-mental terms.[53]

This seems to be an intelligible, materialist goal. Theism reverses the order of explanation.

Theism treats all the laws of nature and the entire cosmos as the result of intentional activity. As such, the account of physical realities and non-conscious, non-mental events, must, in the end, appeal to an intentional reality. One might well reverse all of Dennett's claims within a theistic framework and claim that any explanation of the physical, non-intentional world that did not break out into a deeper, intentional account would be question-begging and not genuinely explanatory. From the vantage point of a fundamentally materialist cosmology, the emergence of consciousness seems strange; it is likened to claiming "then a miracle happens."[54] But from the vantage point of theism, the emergence of consciousness may be seen as something deeply rooted in the very nature of reality. The creation of animal and human consciousness is not some isolated miracle, but a reflection of the underlying structure of reality.

The appeal to a combined case for theism and dualism may be further appreciated in the course of addressing three areas: the nature of divine properties, the nature of mental causation, and the way in which both theism and dualism may be developed along integrative lines. I shall address each briefly.

The nature of divine attributes

Some critics of theism underestimate the resources of theism. Consider, for example, Brian O'Shaughnessy's comment on the problem of matching theism with the precise nature of the cosmos as revealed in modern science:

> Well, four centuries of triumphant advance by the rock-bottom physical sciences of physics cannot but leave some mark on philosophy. When you can predict the wave length of a spectrum line to eight decimal places it is rather more difficult to believe that the underlying reality of everything is spiritual, e.g., an immaterial Deity. After all, should a Deity be so fastidious?[55]

O'Shaughnessy seems to suggest the peculiar position that if there is a God, we should expect the universe to be unspecified, vague or blurred without having any precise features that are measurable to eight decimal places of some length. Alternatively, perhaps if there is a God, God would only make objects that are easily delimited by measurements in whole numbers without any use for fractions. Clearly, the God of classical theism is radically different from O'Shaughnessy's portrait. Traditionally, theists understand God to be unsurpassably great in power, knowledge, and goodness.

Naturalists often fail to adequately appreciate classical Jewish, Christian, and Islamic theistic claims of God's essential aseity or ultimacy. McGinn objects to theists who account for the emergence of consciousness in nature in terms of a broader divine consciousness:

> The theory takes for granted the existence of the conscious agent who is held to explain the existence of all other conscious agents. But if there is a problem about how the conscious beings we see round us come to exist, then there is equally a problem about how the conscious being who creates each of those conscious beings came to exist.[56]

To court theism is to entertain the thesis that there is a being whose properties of omniscience, omnipotence, goodness and *aseity* are not derived from some other agency or explainable in terms of the laws of nature. The theist asks you to consider the possibility that there is such a being and then look at the world to see whether the features of the world make greater or lesser evident sense. Theism is a comprehensive philosophy that offers an explanation of the very existence and continuance of the world and its constituents. I believe that a devalued view of divine attributes leads McGinn and other naturalists underestimate the resources of theism for explaining the very existence of our contingent cosmos. "There is nothing about the intrinsic structure of lumps of mere matter that suggests a divine origin. When we look at rocks through a microscope, no trace of the divine greets the eye."[57] But our eyes are greeted by a cosmos that is contingent (on this point, many naturalists would agree) and it is this feature of the universe along with the emergence of consciousness, that theism can address with great power.[58]

The nature of mental causation

Critics object that dualism leaves us with a complete mystery in terms of mind–body interaction. How can things so remote interact? Invoking theism seems to heighten rather than reduce the problem, for doesn't theism require us also to believe that there can be causal interplay between radically distinct levels of reality? To appeal to theism here would be to explain the obscure in terms of the more obscure.

As R.C. Richardson, B. Mijuskovic and R. Swinburne have forcefully argued, dualist interaction between the mental and physical only looks peculiar if one assumes that the only causal interaction that is intelligible is physical to physical interaction.[59] There is no compelling reason to conclude at the outset that causal interaction cannot be between something physical and something non-physical. There is little consensus today on the nature of physical-to-physical causation, and so little common ground for using an uncontroversial theory of causation in an attack on dualism. Moreover, the assumption that contemporary science somehow rules out such interaction or shows there is none, seems motivated more by philosophy than science. Consider, for example, Churchland's summary of the scientific data against dualism.

> It will be evident from the rest of this book that this familiar hypothesis [dualism]

> is difficult to square with the emerging theory of cognitive processes and with the experimental results from several neurosciences. The doctrine of an immaterial soul looks, to put it frankly, like just another myth, false not just at the edges, but to the core.[60]

But none of the solid scientific findings he cites shows that one *cannot* square dualism with the neurosciences, and it is easy to wonder whether the reason why dualism (or dualism plus theism) looks mythic is because of prior materialist commitments. Given an overall materialist metaphysic, there is not going to be much motivation for recognizing some relatively local exceptions in human and non-human consciousness to the general order of nature.

By way of underlining the way in which Churchland blends his science and metaphysics, I take note of the ease with which a dualist can accommodate many of his scientific claims. Dualists can readily understand Churchland to have located key physical causes in shaping our consciousness. Addressing some particulars, Chruchland writes: "This is where past learning shows itself, where character and insight come in, and where intelligence is ultimately grounded."[61] This may readily be seen as identifying the bodily backing or foundation of our mental lives, but it is quite a different, and decidedly metaphysical, matter to charge that intelligence and so on *is* the brain and body. The blending of science and philosophy emerges at various points. I suggest the following claim seems more metaphysics than science: "Without a neural network in place, there can be no self, neither an emotional self, nor a perceiving self, nor a deliberating self, nor any other kind of self."[62] Imagine science shows the exact correlation in our own case of selves and neural networks and does not give us evidence of there being any immaterial being (God, say) that is a person or person-like. Why would this be taken to demonstrate a *necessary* connection between selves and neural networks? I think Churchland's claim is *informed* by the sciences, but it reflects considerable speculation beyond what is dictated by the sciences alone.

Integrative dualism and theism

Finally, I note that both dualism and theism have been criticized for foisting on us a splintered view of ourselves and the cosmos. Anthony Kenny, P.M.S. Hacker and others propose that dualism offers a huge bifurcation between the person and body such that dualists leave us with a dark skepticism about the mental life of others.

But this also seems like a caricature, especially as dualism is seen in light of theism. In classical theism the world is understood as fashioned by an all-good God who insures the proper functioning of our cognitive, sensory, and volitional powers. The goodness of God is the guarantor of the fit. Given the goodness of God, we might employ Mackie's terms and conclude that a darker, skeptical hypothesis about the

adequacy of our cognitive faculties would be 'queer' (or unlikely).

Some dualists do not emphasize sufficiently the way in which dualism may offer a fully integrated understanding of persons and bodies. The fact, if it is one, that the person and body may be contingently related is no reason to think that in virtually all practical contexts the person and body are functionally identical. Richard Swinburne, for example, depicts a person's relation with their body as follows: "An agent has a body if there is a chunk of matter through which alone he or she can make a difference to the world, and through which alone he or she can learn about the world."[63] As a dualist, I do not deny the claim, but I think it is important to emphasize that in a fully functioning, embodied life, the person and this chunk of matter are functionally indistinguishable. To touch someone's body is not to touch something they are tied to, but, in a perfectly common sense, ordinary sense of the terms, to touch them.[64]

In this last section of this chapter, I have made a plea to understand the philosophy of mind in a broader setting. Since World War II we have seen philosophy of mind expand in important respects; some of this has been motivated by scientific advance and speculation, and some by incorporating other domains in a fuller intellectual environment. For example, much work in philosophy of language has been brought into philosophy of mind, and with great profit, owing to J.L. Austin's seminal contributions. I propose that philosophy of mind may be more fruitfully debated and explored if it too might be incorporated and debated within a broader metaphysical context, one that pares materialism and naturalism over against theism. In this broader setting, we may well see a theistic-dualist philosophy, not as a case of "wallowing" in the supernatural (to use McGinn's term), but as offering a credible, comprehensive philosophy of nature that is profoundly unified, coherent, and able to account for things that otherwise seem queer and anomalous.[65]

Notes

1 J.L. Mackie, *Ethics: Inventing Right and Wrong* (Harmondwsorth, England: Pelican, 1977), 38–42.

2 Ibid., 38.

3 Paul M. Churchland, *The Engine of Reason, the Seat of the Soul* (Cambridge, Mass.: MIT Press, 1995), 211.

4 Jaegwon Kim, *Mind in a Physical World* (Cambridge, Mass.: MIT Press, 1998), 2.

5 William Lycan, *Consciousness and Experience* (Cambridge, Mass.: MIT Press, 1996), 1.

6 Daniel Dennett, *Kinds of Minds* (San Francisco: Basic Books, 1996), 24. See also Dennett, "Living on the Edge," *Inquiry* 36 (1993): 135–59.

7 Nagel suggests a direct link between anti-theism and the growth of "scientism and reductionism of our time" (Thomas Nagel, *The Last Word* [Oxford: Oxford University Press, 1997], 131). See also Georges Rey, *Contemporary Philosophy of Mind* (Oxford: Blackwell, 1997), 23. Churchland refers to dualism in various places as "the religious hypothesis."

8 George Mavrodes, "Religion and the Queerness of Morality," in *Rationality, Religious Belief and Moral Commitment*, ed. Robert Audi and William J. Wainwright (Ithaca, N.Y.: Cornell

University Press, 1986), 213–26.

9 For versions of these arguments, see Galen Strawson, *Mental Reality* (Cambridge, Mass.: MIT Press, 1994); Lynne Rudder Baker, *Explaining Attitudes: A Practical Approach to the Mind* (Cambridge: Cambridge University Press, 1995); William Hasker, *The Emergent Self* (Ithaca, N.Y.: Cornell University Press, 1999); and Victor Reppert, "Eliminativism, Cognitive Suicide, and Begging the Question," *Metaphilosophy 23* (1992): 378–92.

10 Churchland, *The Engine of Reason*, xi.

11 Ibid., 3.

12 Paul M.Churchland, *Metaphysics and the Mind–Body Problem* (Oxford: Clarendon Press, 1979), 64.

13 Churchland, *The Engine of Reason*, 19.

14 Ibid., 8.

15 Ibid., 322.

16 Ibid., 17, 18.

17 Ibid., 324.

18 Ibid., 58 ff.

19 Ibid., 155.

20 Ibid.

21 Ibid., 131.

22 The threat of epiphenomenalism to materialism is lampooned by Dennett in *Brainchildren* (Cambridge, Mass.: MIT Press, 1998), but I think he should be far less sanguine. Kim's *Mind in a Physical World* provides a compelling, sober counter-point.

23 Some materialists note the way in which ordinary terms might be refined under further analysis, scientific and philosophical. Note, for example, William Lycan's comments on "belief":

> The word "belief" . . . points dimly toward a natural kind that we have not fully grasped and that only mature psychology will reveal. I expect that "belief" will turn out to refer to some kind of information-bearing state of a sentient being . . . but the kind of state it refers to may have only a few of the properties usually attributed to beliefs by common sense (William Lycan, *Judgement and Justification* [Cambridge: Cambridge University Press, 1988]).

24 Colin McGinn, *The Mysterious Mind* (New York: Basic Books, 1999), 24.

25 Charles Taliaferro, *Consciousness and the Mind of God* (Cambridge: Cambridge University Press, 1994); and W.D. Hart, *The Engines of the Soul* (Cambridge: Cambridge University Press, 1988).

26 See Taliaferro, *Consciousness and the Mind of God*, chapters two and three for an elaboration and defense of these points.

27 See Frank Jackson, "What Mary Didn't Know," *Journal of Philosophy* 83 (1986): 291–5.

28 Ned Block, "Consciousness," in *A Companion to the Philosophy of Mind*, ed. Samuel Guttenplan (Oxford: Blackwell Publishers, 1994), 213.

29 Rey, *Contemporary Philosophy of Mind*, 57.

30 I defend the reliability of thought experiments and this criterion of identity in "Modal Reliabilism," paper in preparation.

31 Churchland, *The Engine of Reason*, 193.

32 Ibid., 23.

33 Ibid., 201.

34 Ibid., 197.

35 Ibid., 198.

36 See Strawson, *Mental Reality*; and Taliaferro, *Consciousness and the Mind of God*.

37 Lycan, *Consciousness and Experience*, 6–7.
38 Ibid., 48.
39 Ibid., 45.
40 Churchland, *The Engine of Reason*, 192.
41 Ibid., 207.
42 Ibid.
43 Ibid., 197.
44 Ibid.
45 Leopold Studenberg, *Consciousness and Qualia* (Amsterdam: John Benjamins Publishing Co., 1998), 32.
46 Colin McGinn, *The Problem of Consciousness* (Oxford: Blackwell, 1991), xii.
47 See Taliaferro, *Consciousness and the Mind of God*, chapters two and three for a review and analysis.
48 See Hart, *The Engines of the Soul*; J. Harrison, *Essays on Metaphysics and the Theory of Knowledge* (Aldershot, England: Avebury, 1995); Richard Swinburne, *The Evolution of the Soul* (Oxford: Clarendon, 1997); Charles Taliaferro, *Consciousness and the Mind of God*; Taliaferro. "Animals, Brains, and Spirits," *Faith and Philosophy* 12 (1995) ; Taliaferro, "The Perils of Subjectivity," *Inquiry* 40 (1997); Taliaferro, "Possibilities in the Philosophy of Mind," *Philosophy and Phenomenological Research* 57 (1997): 127–37.
49 I defend this in *Consciousness and the Mind of God* as well as in "Mysterious Flames in Philosophy of Mind; Reflections on McGinn's Naturalism," forthcoming in *Philosophia Christi*.
50 Colin McGinn, *The Problem of Consciousness*, (Oxford: Blackwell, 1991), 47.
51 Daniel Dennett, *Brainstorms* (Cambridge: MIT Press, 1978), 83.
52 Daniel Dennett, *Consciousness Explained* (Cambridge: MIT Press, 1991), 454.
53 Rey, *Contemporary Philosophy of Mind*, 21.
54 Dennett, "Living on the Edge," 141.
55 Brian O'Shaughnessy, *The Will*, vol. 1 (Cambridge: Cambridge University Press, 1980), xvii.
56 McGinn, *The Mysterious Mind*, 86.
57 Ibid., 79.
58 See Taliaferro, *Consciousness and the Mind of God*; and Taliaferro, *Contemporary Philosophy of Religion* (Oxford: Blackwell, 1998), for a detailed defense of this claim.
59 See R. C. Richardson, "The 'Scandal' of Cartesian Interactionism," *Mind* 92 (1982): 20–37; B. Mijuskovic, "Brentano's Theory of Consciousness," *Philosophy and Phenomenological Research* 38 (1978): 315–24; and Richard Swinburne, *The Evolution of the Soul*.
60 Churchland, *The Engine of Reason*, 17.
61 Ibid., 11.
62 Ibid., 308.
63 Richard Swinburne, *Providence and the Problem of Evil* (Oxford: Oxford University Press, 1998), 98.
64 See Taliaferro, *Consciousness and the Mind of God*.
65 My thanks for comments on different versions of this paper to the Philosophy Departments at Bethel College, St. Cloud State University, the University of Chicago, New York University, and to Professor Dore Ashton's seminar at Copper Union.

7 Naturalism and libertarian agency

Stewart Goetz

Introduction

David Papineau introduces his book, *Philosophical Naturalism*, with the observation that though the term "philosophical naturalism" ("naturalism", for short) is a familiar one nowadays, there is no universal consensus about its meaning.[1] Papineau believes that naturalism is a commitment to the completeness of physics, where physics is complete in the sense that a purely physical specification of the world, plus physical laws, will always suffice to explain what happens. The concepts of physics, however, change over time. What categories, therefore, will qualify as 'physical' in the ultimate or final physics? We cannot, says Papineau, presently answer this question with any certitude. At best, we can specify one category which will *not* qualify, and that is the category of the psychological which involves propositional attitudes (e.g. beliefs, desires) that represent things as being a certain way (intentionality).[2]

> When I say that a complete physics excludes psychology, and that psychological antecedents are therefore never needed to explain physical effects, the emphasis here is on "needed". I am quite happy to allow that psychological categories *can* be used to explain physical effects, as when I tell you that my arm rose because I wanted to lift it. My claim is only that in all such cases an alternative specification of a sufficient antecedent, which does not mention psychological categories, will also be available.[3]

David Armstrong is another philosopher who addresses the issue of naturalism.[4] According to him, naturalism is "the doctrine that reality consists of nothing but a single all-embracing spatio-temporal system."[5] Contemporary materialism is a form of naturalism and maintains that the single all-embracing spatio-temporal system contains nothing but the entities recognized by physics. Irreducible purpose or teleology has no place in this (or any other) spatio-temporal system as an explanatory

principle because it entails the characteristic of irreducible intentionality, and irreducible intentionality implies the falsity of naturalism. Thus, Armstrong says, "I suppose that if the principles involved [in analyzing the single, all-embracing spatio-temporal system which is reality] were completely different from the current principles of physics, in particular if they involved appeal to mental entities, such as purposes, we might then count the analysis as a falsification of Naturalism."[6]

If we take our lead from Papineau and Armstrong, naturalists are those who maintain that the ultimate *explanation* of things will not include psychological (intentional, mental) categories and, therefore, will exclude teleology. What, then, is the ontological status of the psychological according to naturalism? Is is not even real? 'Hard' naturalists not only deny its relevance for ultimate explanation, but also its reality.[7] 'Soft' naturalists, while denying the psychological's ultimate explantory relevance, try to preserve its reality. Such naturalists often invoke the concept of supervenience and maintain that the psychological supervenes on the physical. Thus, David Chalmers argues for naturalism and maintains that while "experience is superfluous in the *explanation* of behavior,"[8] the fundamental theory of the world will include psychophysical supervenience laws which tell us how experience arises from physical processes.[9] John Heil, who is also a naturalist, says the following about naturalism: "Naturalism in the philosophy of mind is the view that mental characteristics are determined by or *supervene* on features of agents comprehended by the natural sciences."[10]

A common thread running through all conceptions of naturalism, then, is that the fundamental explanation of any event is non-psychological in nature. Naturalism, therefore, ultimately implies the falsity of teleological explanation. Because naturalism excludes teleology, it also excludes libertarian freedom or agency. According to libertarianism, a choice is an undetermined mental action which is explained teleologically in terms of a purpose or goal of its agent. A teleological explanation of a choice, because it makes reference to the concept of a goal, end, or purpose derived from the representational content provided by a propositional attitude, is believed by the naturalist to be either eliminable for, reducible to, or dependent upon a fundamental non-teleological explanation of that event.

In addition to rejecting the reality of libertarian freedom, naturalists also deny substance dualism (dualism, for short), which is the view that there is a non-physical soul or mind which has irreducible psychological properties.[11] Given the naturalist's rejection of both libertarianism and dualism, an interesting and important question is whether the truth of the former implies the truth of the latter. If it does, it would not be surprising to find that naturalists reject libertarianism because they reject dualism.

The course of my argument in this paper is as follows. In Section I, I set forth a non-causal account of libertarian agency in which a choice is an essentially uncaused mental action explained by a *telos* or purpose of its agent, and I make clear why such

an account is an intuitively plausible conception of free mental action. Surprisingly, but for no good reason, a non-causal account of human freedom is normally rejected by libertarians themselves. Therefore, in Section II I defend my non-causal agency theory against the criticisms of other libertarians as a way of highlighting the centrality of teleological explanation for libertarianism. Section III consists of an explanation of why naturalists believe that teleological explanation of choice implies the truth of dualism and a discussion of the alleged problem of causal interaction which they assert is a decisive reason for rejecting dualism. I claim that dualism is no worse off than any soft naturalist's view of the mind when it comes to explaining the relationship between the psychological and the physical. Thus, if naturalism implies the falsity of dualism and libertarianism entails dualism, if we have libertarian freedom, which one prominent soft naturalist concedes we certainly seem to have,[12] naturalism is false.[13] Finally, in Section IV I briefly examine an emergentist alternative to the view that libertarianism implies dualism, and explain why it is inadequate.

I

A desideratum which any libertarian theory must satisfy is that of being able to explain how it is that mental actions such as choices differ ontologically from mental non-actions or mere happenings. One way to account for this is in terms of the extrinsic relationship of causal ancestry: mental actions differ from mental non-actions in terms of their different causal histories. This is the approach taken by libertarians who espouse agent-causation. A more intuitively satisfying view is to say that mental actions differ intrinsically from mental non-actions. Because a choice is a mental action, how can it be intrinsically active? In the following way: most generally, a mental *power* is an ontologically irreducible property which is exhibited by an entity. Corresponding to a mental power is the *exercising* of that power which is an event. Exercising a mental power is *acting*, and mental action is essentially *uncaused*. One kind of mental power an agent possesses is the mental power to choose. When an agent exercises it, he chooses. Thus, (i) a choice is the exercising by an agent of his mental power to choose, where (ii) the exercising of the power to choose is essentially an uncaused event.

Consider (i). Carl Ginet has recently defended an account of mental action which maintains that any mental act differs intrinsically from passive mental events. According to him, a mental act's intrinsic active nature consists in its having an 'actish phenomenal quality'. This actish phenomenal quality lacks the complex structure of a causal relation and, by itself, is enough to make a mental event a mental action.[14] I believe it is a mistake, however, to characterize the intrinsic active nature of a mental action as a phenomenal quality.[15] Such a characterization suggests that a mental act has a distinctive *quale* or feel about it which makes it intrinsically active and distinguishes it from a passive mental event. The exercising of the power to choose,

however, like any mental act, has no intrinsic feel or *quale* about it. Thus, while I may feel tired after exercising my power to think, which I do when I think for a long time about the issue of free will, this active thinking itself has no intrinsic *quale* any more than choosing does. The active nature of a mental event consists solely in its being the exercising of a mental power.[16]

If a mental action is an event which is the exercising of a power, what is a mental non-action or mere happening? It is what I will call the *actualization* of a mental *liability*. Like a mental power, a mental liability is an ontologically irreducible property which is exhibited by a subject. When a subject's mental liability is actualized, a *caused* event occurs with respect to which he is a patient. One mental liability a subject has is the liability to believe. When his liability to believe is actualized, he believes a proposition and he is a patient with respect to believing that proposition.

In summary, there are two types of mental properties, namely, powers and liabilities. These two kinds of properties are inherently different from each other and each is an ultimate category in our ontology. Corresponding to these two kinds of mental properties are two kinds of events, namely, the exercising of a mental power (an action) and the actualization of a mental liability (a passion). Like the properties themselves, these two kinds of events are inherently different from each other such that any exercising of a mental power is intrinsically distinguished from any actualization of a mental liability.

Support for (ii) is conceptual in nature. An event which is efficiently caused is produced by that cause and as such is an occurrence with respect to which its subject is essentially passive. An event is being made to occur to the subject and it (the subject) is not active with respect to that event. Because an exercise of mental power is active in nature, it is not produced and, thus, cannot be caused.

The intrinsically active and uncaused nature of choosing (ontology) is confirmed by the epistemology of action. In commenting on causal theories of action, Harry Frankfurt makes the following point about an agent's knowledge of his action:

> They [causal theories] are therefore committed to supposing that a person who knows he is in the midst of performing an action cannot have derived this knowledge from any awareness of what is currently happening but that he must have derived it instead from his understanding of how what is happening was caused to happen by certain earlier conditions. . . . This is what makes causal theories implausible. They direct attention exclusively away from the events whose natures are at issue, and away from the times at which they occur.[17]

Frankfurt's remarks suggest the following point about mental action such as choice: It is an epistemological feature of an agent who knows that he is making a choice that he knows this while he is choosing. Given this fact, it is natural to think that the agent

knows that he is choosing by being aware of the choice which he is making. However, if he knows in this way that he is choosing, it seems to follow that choosing is intrinsically different from a mere happening or passive event and that he is aware of this difference. On a causal theory of action, however, an agent who knows that he is choosing cannot possess this knowledge in virtue of his awareness of the choice itself. This is because on a causal theory of action a choice is not intrinsically different from a mere happening. In themselves, the two are indistinguishable. Therefore, an agent can know that he is choosing only by being aware of causal differences which distinguish the two events.

According to Frankfurt, a causal account of action implies that the agent's attention must be directed away from the mental action which he is performing in order for him to know that he is acting. Thus, a causal theory is unable to account for a significant epistemological feature of mental action. However, while Frankfurt has directed our attention to an important epistemological feature of mental action, he unnecessarily obscures his point by suggesting that the problem with a causal theory of action stems from what it implies about the temporal distance separating a mental action from its cause. Thus, Frankfurt not only maintains that a causal theory entails that an agent who knows that he is choosing must derive his knowledge from an awareness of what causes his choice, but also he maintains that the cause is a certain *earlier* event. Because it is, the agent who knows that he is choosing must be directing his attention away from the *time* at which he is choosing.

If the problem with a causal theory of action pointed out by Frankfurt were essentially linked with this temporal issue, the causal theorist would have a rather obvious response. He could simply stipulate that the cause which distinguishes his choice from a passive event happens simultaneously with the choice.[18] Because the cause occurs simultaneously with the choice, there is no epistemological problem of the kind noted by Frankfurt. It is true that the agent who knows that he is choosing must look to a causal antecedent, but the causal antecedent is, in virtue of its simultaneity with its effect, known immediately by the agent.

The epistemological problem with a causal theory of action pointed out by Frankfurt is not essentially linked with temporal considerations. Rather, it arises in light of the fact that agents who know that they are performing mental actions possess this knowledge by being aware of the actions themselves, without reference to any cause of them. In the case of a choice, the exercising of the power to choose by the agent is essentially uncaused and intrinsically active and the agent knows that he is choosing by simply being aware of the choice (the exercising of the power to choose) itself.

If agents are directly aware of choosing and choices are essentially uncaused events, it seems to follow that agents are directly aware of the lack of causation of their choices. But is this plausible? Some critics of libertarianism maintain that it is thoroughly implausible "because undetermined choice (as opposed to our lack of

recognition of causes for our choices) is not introspectible at all."[19] We cannot introspect indeterminacy because we cannot "observe the lack of causes."[20]

If an agent had to conduct a search for the lack of causes, it would be implausible to maintain that he is aware that his choice is uncaused because he is aware of the lack of any causes which produce his choice. It would always be the case that he might have overlooked a cause which is actually there. It does not follow from this point, however, that an agent can at best only fail to observe the presence of any causes. It is *a conceptual truth* that a choice is uncaused because a choice is the exercising of the power to choose and exercisings of mental powers are essentially uncaused. Thus, because an agent can be directly aware of exercising his power to choose, he knows that his choice is uncaused because this is entailed by the nature of what it is to be a choice.

If a choice is uncaused, what explains its occurrence? Because a choice is a mental *action*, the most plausible and intuitively satisfying answer is that a choice, like any action, is explained in terms of a reason, where a reason is a goal, end, or purpose for which an agent chooses. The concept of acting for a reason is correctly captured or expressed with the "in order to" locution. Thus, we say that Kathryn chose to be an exchange student in Germany in order to learn to speak the German language. To explain a choice by a reason is to explain it teleologically, not causally.

An agent's reason for choosing is grounded in the representative, semantic contents of certain of his psychological attitudes. These psychological attitudes are his desires and beliefs. For example, when Kathryn chose to be an exchange student in Germany, she had a desire to learn the German language and believed that spending several months living with a family in Germany and attending Hochschule would achieve the goal of learning German. The desire that she learn German provided her with her purpose or goal expressed by "in order to learn German" which explained her choice to be an exchange student. Thus, the semantic content of Kathryn's desire was not identical with her reason for choosing, but provided the basis for the reason for her choice. Similarly, the event of her coming to have or the event or state of her continuing to have the relevant desire did not cause her to make her choice, as is typically maintained by causal theorists.[21]

II

I have developed an account of libertarian freedoom (free will) where indeterministic choices are uncaused events which are adequately explained in terms of teleology alone. Not all libertarians have endorsed such a non-causal account. Indeed, most have been reluctant to do so.[22] In this section, I will set forth and respond to five libertarians who have claimed that teleological explanation of choice must be supplemented by causal explanation.

To begin, consider Richard Taylor's criticism of non-causal libertarianism.

According to Taylor, "[o]nly the slightest consideration will show that this simple denial of determinism has not the slightest plausibility."[23] Why not? Because, while non-causal libertarianism avoids picturing an agent as a puppet, it does so

> by substituting something even less like a human being; for the conception that now emerges is not that of a free person, but of an erratic and jerking phantom, without rhyme or reason at all. . . . There will never be any point in asking why [choices] occur, or in seeking any explanation of them, for under the conditions assumed there is no explanation. They just happen, from no causes at all.[24]

From the fact that a choice is uncaused, however, it does not follow that it has no explanation. To conclude this is to identify causation with explanation. Contrary to what Taylor's argument assumes, causation is only one kind of explanation. Causation stands to explanation as species to genus. Teleology is another form of explanation. Moreover, Taylor's rejection of non-causal libertarianism is puzzling because he recognizes that "[i]n the case of an action that is both free and rational, it must be such that the agent who performed it did so for some reason, but this reason cannot have been the cause of it."[25]

Next, consider Robert Kane's view of free will. He asserts that in formulating their accounts of free will, libertarians should not appeal to special kinds of entities or causes (substances, properties, relations, events, states, etc.) which are not needed by non-libertarian accounts of free will.

> The only difference allowed between libertarian and nonlibertarian accounts is the difference one might expect – that some of the events or processes involved in libertarian free agency will be indeterminate or undetermined events or processes. But these undetermined events or processes will not otherwise be of categories or ontological kinds that do not also play roles in nonlibertarian accounts of free agency . . . – the difference being that in nonlibertarian theories, these events or processes need not be undetermined. Such differences as there are between libertarian and nonlibertarian theories should flow from this difference alone, and the task will be to make sense of a libertarian freedom satisfying the plurality conditions, given this difference.[26]

The plurality conditions to which Kane refers in the last-quoted sentence consist of an agent's having reasons for alternative courses of action; there being no compulsion or coercion of the choice which is made by him for the relevant reasons; and his having control over that choice in the sense that *he* makes it for those reasons.[27]

It is clear that the non-causal account of free will I have developed in which a choice

is an uncaused exercising of the power to choose does not introduce or appeal to categories or kinds of entities beyond those mentioned by Kane. An uncaused choice is an exercising (an *event*) of a power (a *property*) which is causally undetermined. Moreover, my account of free will satisfies the plurality conditions set forth by Kane. Nevertheless, he criticizes the libertarian position which maintains that only reasons-explanations can be given of choices because

> it stands in the way of understanding the place of free will in the natural order in which causal explanations of deterministic or probabilistic kinds hold sway. Causal explanations cannot be the whole story of free will, but they must be part of the story if free will is to have a place in the natural order. And since deterministic causal explanations are ruled out for undertermined free willings, it is important to ask how explanations of them in terms of reasons are related to probabilistic or nondeterministic causal explanations.[28]

Kane asserts that the choice which is made by an agent must be probabilistically and, thereby, indeterministically caused by the reasons for making that choice,[29] where both the choice[30] and the reasons[31] for which it is made have corresponding neural connections or brain-wave patterns. But what do indeterministic, probabilistic causes explain about the making of a choice that is not teleologically accounted for by the reasons for making it?[32] Kane's view seems to be that causation involving an agent's reasons is required as a link to explain why the choice is "the *agent's* doing rather than a mere happening".[33] Three responses are in order.

First, if Kane is claiming that causation is required to make the choice an action, he is simply mistaken. As I explained in setting forth my account of the nature of choice in the previous section, the view that a choice is intrinsically active is most plausible in light of both libertarian and epistemological considerations about action.

Second, if Kane is claiming that causation is necessary to account for why a choice is *mine*, as oppposed to some other agent's, he is again mistaken. As I also asserted in Section I, the view that a choice is an exercising of the mental power to choose, and that a mental power is an ultimate and irreducible property of an agent, is plausible in its own right. In addition, it is supported by the view that the essences of physical particles are constituted, at least in part, by (1) their causal powers to attract and repel other particles with similar powers and (2) their liability to be moved. Moreover, one would want to know from Kane what makes the reasons which, on his view, probabilistically cause my choice, *my* reasons as opposed to some other agent's. If Kane claims that these reasons just are my reasons (a view which seems as eminently plausible as that about choice which he rejects), then it is incumbent upon him to explain why it is that the reasons which probabilistically cause a choice can just be an agent's, but the choice made for those reasons cannot. Some property of an agent

must essentially characterize it, and there is no better candidate for such a property than the power to choose (though the power to choose need not be the only such property).

Finally, I believe it is plausible to maintain that probabilistic causation is not a form of *explanation* at all. It is no more than a *description* of how often an event of a particular kind occurs in light of its explanation. Kane, himself, seems to realize this. Thus, he states that "while the probabilistic explanation tells us which outcome is more likely to occur, it does not tell us why one occurs rather than the other *in the particular case*; and this is true whether the outcome that occurs in the particular case is more probable or less probable."[34] Kane also acknowledges that there is a teleological explanation of a particular choice.[35] Thus, what is more or less probable is whether a particular choice will be made by an agent for a relevant reason in light of his past record of having or not having made choices of the same kind in similar situations.

Randolph Clarke is a third libertarian who maintains that a non-causal libertarianism cannot adequately account for free will (choice). According to him, free will is a certain kind of agent-control, where agent-control is fundamentally a matter of what does and does not causally produce a free act of will.[36] Clarke notes that it is possible to maintain that agent-control is a matter of the complete absence of causation of an action. Like most other libertarians, however, he thinks a non-causal libertarianism is inadequate[37] and causation in the form of agent-causation must be invoked to account for free will. As best as I can ascertain, Clarke's reason for believing that a non-causal libertarianism is inadequate is an argument put forth by Donald Davidson that reasons-explanation must be understood as a form of causal explanation. Clarke maintains that no non-causal theory has met the challenge of this objection.[38] Therefore, in what follows, I will summarize and respond to Davidson's argument.

According to Davidson, if reasons-explanation is not a form of causal explanation, then there is no way to account for the distinction between having a reason to choose and choosing *with* it, and having a reason to choose and choosing *because* of it. One may *justify* a choice by citing a reason one has even if in choosing one did not choose because of it, but one cannot *explain* a choice by citing a reason unless one chooses because of it.[39]

To illustrate this objection, consider an executioner named Smith. Smith is ordered to put to death a certain individual, Charles. Charles was caught in the act of, confessed to, and was convicted for murdering a woman who happened to be Smith's wife. Smith has two reasons for putting Charles to death. On the one hand, he believes that it is his job and moral duty to put to death persons guilty of the heinous crime of murder. On the other hand, Smith desires to get revenge against Charles by killing him. Smith has more than one reason to kill Charles. Both reasons justify the act. After putting Charles to death, Smith tells a reporter that he chose to put Charles to death

in order to carry out what he believed was his job and moral duty, and he did not choose to kill him in order to get revenge. According to Davidson, this can only mean that Smith's having the one reason caused him to make the choice and his having the other reason did not.

On a teleological account of how reasons explain an action, Smith chose to put Charles to death in order to carry out what he believed was his job and moral duty and did not put Charles to death in order to get revenge. In saying that Smith chose in order to do the former but did not choose in order to do the latter, it is the case that the former was an end for which he chose and the latter was not. The former's being an end for which Smith chose is no less objective than its being the case that one event was the cause of an effect and another was not.[40]

A fourth libertarian who maintains that teleological explanation alone is not adequate to explain libertarian choice is Timothy O'Connor.[41] He maintains that there must be causation to link the reason for making a choice with the choice, because he, too, thinks that teleological explanation alone cannot account for the distinction between having a reason and choosing with it and having a reason and choosing because of it. According to O'Connor, the requisite causation is agent-causal in form. He asserts that an agent has the causal capacity to cause the coming to be of an intention to act, where the coming to be of an intention is an event-part of a choice which is made for a reason. He claims that the causation by the agent of the coming to be of an intention has no cause and is "dependent upon the reason he has . . . for acting [choosing] in that way. . . . For the agent's free exercise of his causal capacity provides a necessary link between reason and action, without which the reason could not in any significant way explain the action."[42]

Contrary to what O'Connor believes, if there is a problem with explaining how a reason explains an uncaused choice, there is a problem with explaining how a reason can explain an agent's uncaused exercising of his power to cause the coming to be of an intention, where the exercising of this power depends upon the reason for the choice of which the coming to be of the intention is a part. Because there is no problem for the non-causal libertarian view which maintains that a reason alone can explain an uncaused choice, agent-causation is superfluous to an adequate account of freedom.

The last libertarian who contests a non-causal libertarianism of the kind I advocate is Ginet.[43] Like Clarke and O'Connor, Ginet believes that there must be some way to account for the distinction between having a reason and acting with it and having a reason and acting because of it. Ginet maintains that with a typical bodily action done for a reason, such as opening a window in order to let in fresh air, an agent S intends of that action that by it she will let in fresh air. The intention explains the action in virtue of its content referring directly to the action. In this example, the content of the intention is 'by *this* opening of the window I will let in fresh air'. Moreover, it is not

required of the intention that it play a causal role in the opening of the window. It is possible for S not only to have a reason to open the window but also to have a reason for keeping the window closed (e.g. in order to stay warm). In this case, the agent acts *because* of the first reason and only *with* the second in virtue of the specified intention. The intention allegedly forges a link between the reason for which the action is done and the action itself. To illustrate Ginet's position, consider an example of his where *S* urgently needs her glasses which she has left in a person, *R*'s, room where *R* is now sleeping.

> *S* has some desire to wake *R*, because she would then have *R*'s company, but also some desire not to wake *R*, because she knows that *R* needs the sleep. *S* decides to enter *R*'s room in order to get her glasses, knowing as she does so that her action will satisfy her desire to wake *R*. Could it nevertheless be true that *S* did not intend of her action that it wake *R*? . . . It seems right to say that *S* did not intend to wake *R* if *S* was so disposed that, had it turned out that her entering the room did not wake *R*, *S* would not have felt that her plan had failed to be completely realized, and she must then either wake *R* in some other way or decide to abandon part of her plan. And *S*'s being thus uncommitted to waking *R* is quite compatible with *S*'s expecting and desiring to wake *R*.[44]

Before evaluating Ginet's position, it is helpful to clarify how his libertarian account of freedom differs from my own. On my view, freedom pertains to an agent's choice, and while the concept of intention provides an explanatory connection between the choice and the chosen action, it provides no such connection between the reason and the choice itself because the concept of choosing for a reason is primitive in nature. Thus, *S* only intends the action of her getting the glasses for a reason because she first chooses to perform that action for that reason. Intention provides no explanatory link between reasons and the choices made for those reasons.

At this point, Ginet might argue that it is false to maintain that free agents only intend to perform those actions which they first choose to do. Free agents also intend their choices to act. They must intend those choices because their choices are teleologically explained by reasons and an intention is required to forge a link between those choices and their reasons, if the libertarian is to present an adequate account of the distinction between acting with a reason and acting because of it.

In response, one needs to ask 'How do free agents intend their choices?' What *explains* such intentions? The reasons *R* for which the choices are made? But if *R* can explain these intentions directly without the need for something to forge an explanatory link between the two of them, why cannot *R* explain the choices directly without the need for intermediate intentions to link the choices with their reasons? Do further reasons *R'* explain the intentions? But if free agents intend their choices

for *R′*, will they not need second-order intentions to link the first-order intentions with their reasons *R*? Here, a vicious regress seems unavoidable. The only way to avoid it is to say that free agents intend their choices for reasons *R′*, and that is the end of the matter. But if there is nothing problematic with saying this, there is no problem with saying that a free agent such as *S* chooses for reasons *R*, and that is the end of the matter. Thus, by maintaining that agents intend their choices, Ginet has not really explained the distinction between acting with a reason and acting because of it. Rather, he has merely relocated the place where the distinction needs to be accounted for.

Ginet might respond to this criticism of his view by maintaining that it is preferable to talk in terms of the act of adopting the intention to perform act *A*, instead of the act of choosing to *A*,[45] where the act of adopting the intention to *A* forges an explanatory link between *A* and the reason *R* for which *A* is done. Even if we concede to Ginet this preference, it can still be shown that intention cannot plausibly be used to account for the distinction between acting with a reason and acting because of it. This is because adopting an intention is itself an act and, as such, is explained by a reason. What forges an explanatory link between the adopting of an intention and the reason for which it is adopted? Is there some further intention which is adopted? If there is, then it will have to be adopted for a reason and an explanatory link will have to be forged between it and that reason. A vicious regress like that described in the previous paragraph is avoidable only by saying that an intention can be adopted for a reason, without any explanatory relation between it and that reason being forged by a further intention. But if this is the case, then the explanatory connection between a reason and the act of adopting an intention is ultimately primitive, and intention cannot be used in the way Ginet suggests to distinguish between having a reason and acting with it and having a reason and acting because of it.[46]

III

Classically, libertarianism was held to be incompatible with naturalism because the former was indeterministic in nature while the latter was deterministic. The determinism espoused by those who can be called 'classical naturalists' was causal in nature and was often expressed by saying that the causal laws of nature and the past together entail the occurrence of one course of events. Thus, causal determinism was, one might say, past-to-present in form. There is, however, a different kind of determinism which is currently popular among naturalists. According to contemporary, non-classical naturalists, even if classical past-to-present causal determinism is false and what is physical is ultimately indeterministic (as some understandings of quantum physics maintain) in nature, it is still possible and plausible to believe that determinism is true because it is reasonable to believe in the

general ontological thesis that reality is a multilayered hierarchy which consists of levels of entities with their characteristic properties.[47] It is assumed by present-day naturalists that there is a lowest, fundamental, or bottom level of reality which consists of what micro-physics considers to be the most basic particles out of which everything is composed. On 'top' of the lowest level are higher, intermediate-level entities (e.g. chemical, biological) with their distinctive properties. Mental properties are properties of human beings (brains or central nervous systems) which are higher (highest?)-level, macro-entities.[48] There is a *dependency* relation between the lower-level, physical properties of micro-objects and the higher-level mental properties of human beings such that no human being can have its mental properties unless it has physical properties, and the lower-level physical properties *determine* the higher-level mental properties in the sense that nothing can be just like a given human being in all of its physical properties without its also being just like it as regards its mental properties. That is, physical indiscernibility entails psychological indiscernibility.[49] Contemporary bottom-to-top determinism is typically characterized by its naturalist proponents as the supervenience of the mental on the subvenient physical, and, thus, might be termed 'supervenience determinism'. In supervenience determinism, the supervenience of the mental on the physical is held to be *asymmetric* in character in the sense that the higher-level mental properties are dependent upon and determined by the lower-level physical properties, but the bottom-level physical properties are not dependent upon and determined by the higher-level mental properties.[50] In summary,

> [a]ccording to some philosophers, mind–body supervenience gives us the right kind of physicalism: It respects the primacy of the physical by giving a clear sense to the idea that the physical determines the mental. Without the instantiations of appropriate physical properties, no mental property can be instantiated, and what particular mental properties are instantiated depends wholly on what physical properties happen to be instantiated.[51]

Thus, on what I will call the contemporary 'supervenience naturalist's' view of the world (supervenience naturalists endorse supervenience determinism), mental properties and events cannot 'float free' of physical properties and events. Because one's physical life superveniently determines one's mental life, one's mental life is not autonomous from one's physical life.

The implications of the bottom-to-top determinism of the mental by the physical in the supervenience naturalist's view of the world for free will has not been widely discussed.[52] Searle is one supervenience naturalist who has raised and addressed the relevant implications, and in what immediately follows I will summarize his discussion of them. His comments are especially helpful because they help to tease out the logical

relationships between free will, dualism, and naturalism. What I ultimately hope to show in this section is that there is good reason to think that libertarianism implies the falsity of naturalism for a more basic reason than its incompatibility with determinism. This reason is that libertarianism implies dualism, and dualism implies the falsity of naturalism.

According to Searle, the libertarian is right about this much: belief in libertarian freedom is not the invention of philosophers.[53] On the contrary, it is something which all of us believe as ordinary people in everyday life:

> [I]f there is any fact of experience that we are all familiar with, it's the simple fact that our own choices . . . seem to make a difference to our actual behaviour. . . . We know we could have done something else, because we chose one thing for certain reasons. But we were aware that there were also reasons for choosing something else, and indeed, we might have . . . chosen that something else. . . . Human freedom is just a fact of experience.[54]

Though it is obvious to us from our experience that we seem to have free will, Searle claims that modern science poses problems for its existence. It does so because of its bottom-up view of the world in which nature consists of particles and their relations with each other, and everything about the world can be accounted for in terms of these particles and their relations.

> Why exactly is there no room for the freedom of the will on the contemporary scientific view? Our basic explanatory mechanisms in physics work from the bottom up. That is to say, we explain the behaviour of surface features of a phenomenon such as the transparency of glass or the liquidity of water, in terms of the behaviour of microparticles such as molecules. And the relation of the mind to the brain is an example of such a relation. Mental features are caused by, and realised in neurophysiological phenomena . . . .[55]

Searle maintains that his account of the mind–brain relationship is a form of supervenience naturalism.

> On the account that I have been proposing, mental states are supervenient on neurophysiological states in the following respect: Type-identical neurophysiological causes would have type-identical mentalistic effects. . . . On this characterization of the supervenience relation, the supervenience of the mental on the physical is marked by the fact that physical states are causally sufficient . . . for the corresponding mental states. . . . [S]ameness of

> neurophysiology guarantees sameness of mentality . . . .[56]

Given the supervenience naturalist's layered picture of the world and the place of the mental in it as a higher-level macro-feature which is supervenient upon the lower-level micro-features, Searle claims the following would have to be the case if we were to have free will:

> In order for us to have radical [libertarian] freedom, it looks as if we would have to postulate that inside each of us [our physical bodies] was a *self* that was capable of interfering with the causal order of nature. That is, it looks as if we would have to contain some *entity* that was capable of making molecules swerve from their paths. I don't know if such a view is even intelligible, but it's certainly not consistent with what we know about how the world works from physics.[57]

As Searle describes matters, the problem for free will is that it requires that a dualist self or mind be able to causally affect the micro-physical world and such causal interaction is problematic, if not unintelligible. It is relevant to note at this juncture, however, that if such causal interaction is a problem for free will, it is not one which is distinctively generated by the thesis that the mental is superveniently determined by the physical. It is just an application of the classic objection to dualism in the context of supervenience naturalism. Therefore, before addressing the alleged problem of mental-to-physical causation for dualism, it is important to understand why it is that supervenience naturalists believe that libertarianism implies the truth of dualism. In order to acquire this understanding, it will be helpful to focus on the *explanatory* nature of the determinism in the supervenience relation itself.

As I indicated at the outset of this section, classically, the explanatory character of naturalism consisted of causal determinism. Supervenience naturalists, however, say little, if anything positive, about the explanatory nature of the determinism in the supervenience relation, except that it is deterministic. Indeed, they seem to know more about what the explanatory nature of the determinism is not than about what it is. To see that this is the case, consider Searle's view. In a spirit of continuity with the classical naturalist view, he maintains (see the penultimate quote) that physical states causally determine their supervening mental states. According to Searle, supervenience determinism is a form of causal determinism. He is careful to point out, however, that the causation in the supervenience relation is not event-event in kind.[58] This will strike many, if not most, as very puzzling or even incoherent, because causation seems to be essentially a relation involving *events* of objects or substances. For example, in terms of the ontology which I set forth in Section I, causation essentially involves the eventful exercising of a causal power by an agent which

produces the eventful actualization of a liability in a patient.

Kim is another supervenience naturalist, and though he denies that determinism in the supervenience relation is a form of causal determinism of any kind (whether event or not-event),[59] he has little, if anything, to say in a positive vein about it. Heil summarizes the inabilility of supervenience naturalists such as Kim to say anything positive about the explanatory nature of the determinism in the supervenience relation in the following way:

> Although modal notions constitute an important ingredient of the concept of supervenience, they do not exhaust that concept. . . . [W]e might inquire as to the remaining aspects. These include, first, a notion of dependence, according to which a property or characteristic in the supervenient family requires for its instantiation the instantiation of a property or characteristic in the subvenient family; and second, a notion of determination, according to which supervenient properties or characteristics are instantiated *in virtue of* or *because of the* instantiation of some subvenient property or characteristic. . . . Offered as an analysis of supervenience, these remarks would be excessively thin. We should want to know what dependence and determination amounted to, and how they bear on explanation. . . . This difficulty is . . . one about which I have nothing exciting or original to suggest.[60]

How about what the explanatory nature of determinism in the supervenience relation is not? For present purposes, it is crucial to note one thing which supervenience naturalists (in virtue of their *naturalism*) maintain it is not, and that is that it is not teleological in nature. Whatever the explanatory nature of the determinism in the supervenience relation is, supervenience naturalists are united in affirming that what explains a choice (provided a supervenience naturalist acknowledges the reality of a choice as a genuine mental event) is not a purpose or goal of an agent (a teleological explanation). The bottom-to-top layered picture of the world where the mental is a higher-level feature excludes future-to-present explanation of a choice in terms of a goal or purpose. Thus, while supervenience naturalism is incompatible with libertarianism because of the determinism present in the supervenience relation, at a more basic level it is incompatible with libertarianism because like all forms of naturalism[61] it denies any explanatory role for teleology.

Supervenience naturalists are committed to the thesis that there are no undetermined choices which are made by their agents for purposes. To clarify why they exclude teleological explanation, it is helpful to return to one of Searle's comments quoted earlier. As he analyses the issue of human freedom, if we were to concede the reality of libertarian choices made for purposes, then we would be committed to the view that mental events, rather than being dependent surface

features of the world which are determined to occur by deep subvenient micro-physical characteristics, would themselves be independent deep features of the world. Searle believes (as do most, if not all, supervenience naturalists) that in order for a mental event such as a choice to occur and be explained teleologically, it would have to be an independent deep occurrence of the world, and this would imply that a *self* exists which is an independent deep entity that makes the choice for a purpose. And if this self exists and is able to choose for a purpose, then if it is to realize or accomplish its purpose for choosing, it will have to be able to causally affect what goes on at the micro-physical level of the world. In short, in order for libertarian freedom to be real, a teleological explanation must be the ultimate explanation of a choice. And in order for a teleological explanation to be the ultimate explanation of a choice, a substantial self (mind or soul) which makes a choice for a purpose must exist at the deepest level of reality and be able to causally interact with the ultimate or deep entities of the physical world. Stated in slightly different terms, if libertarianism is true, there is a form of explanation, namely, teleological, which is equally fundamental as the causal explanation operative at the micro-physical level of reality, and on those occasions when an agent chooses for a reason, events at the micro-physical level of the world occur only because a soul causes them to occur in accomplishing its goals. Because goals or purposes must be grounded in the representational character of propositional attitudes (mental contents) of minds, if there is teleological explanation at the most basic level of reality, then it follows that minds with their representational contents exist at the most basic level of reality. Therefore, if choices are ultimately explained teleologically, the dependency of the mental on the physical is undermined. In order for us to have free will, dualism must be true. From the perspective of a supervenience naturalist such as Searle, if we accept the reality of free will, we fall into the "absurdity of supposing that there are two kinds of substances . . . in the world," and "[w]e are reluctant to concede any of the commonsense facts that sound 'Cartesian,' because it seems that if we accept the facts, we will have to accept the whole of Cartesian metaphysics."[62] Kim echoes Searle on this issue:

> [T]o abandon the physical causal closure is to retrogress to the Cartesian picture that does not allow, even in principle, a complete and comprehensive physical theory of the physical world. On the Cartesian dualist model, any theory that gives full coverage of the physical would have to invoke nonphysical causal agents. This is something that no serious physicalist will find palatable.[63]

The thinking of the supervenience naturalist to this point can be summarized as follows: free will is just a fact of experience, but free will requires the truth of dualism. There is, however, the problem of causal interaction between an immaterial soul and a micro-physical entity which implies the falisty of dualism. Given the falsity of dualism,

libertarianism is false. Thus, contrary to the way things seem, there is no free will.

It is now time to examine the alleged problem of causal interaction. To begin, it is appropriate to spend a few moments on the issue of the intuitive or initial plausibility of dualism. Is the soul, like the free will which implies its existence, just a fact of experience? It has seemed so to many, if not most, people. Among philosophers, Descartes noted in the Sixth Meditation that "as a matter of fact, when I consider the mind, that is to say, myself inasmuch as I am only a thinking thing, I cannot distinguish in myself any parts, but apprehend myself to be clearly one and entire . . . ."[64] In other words, on the basis of self-awareness, it seemed to Descartes that he was a substantial mind which lacked proper parts (a simple entity). That this 'simple view' of the self is not an aberration is recognized by several contemporary non-dualist philosophers of mind. For example, Nagel states that the self exhibits an "apparent unique simplicity and indivisibility,"[65] while David Braine captures the seeming simplicity of the self in the following words:

> [T]here is no question of dividing an 'I' . . . into parts. The 'I' . . . presents itself as undivided and indivisible. . . . [W]hen we think of the 'I' . . . as undivided and indivisible, we are *somehow* denying that, as such, it has parts at all rather than thinking of some special kind of integratedness of its parts.[66]

It is important to stress that these philosophers are noting the features of first-person self-awareness which explain why it is that the ordinary person is a dualist. They are not foisting a philosophical invention on the ordinary mind. Thus, Searle notes that "when I lectured on the mind-body problem in India", [I] "was assured by several members of my audience that my views must be mistaken, because they personally had existed in their earlier lives as frogs or elephants, etc. . . ."[67] Searle goes on to add that "given what I know about how the world works, I could not regard their views as serious candidates for truth."[68]

As we have already seen, what Searle thinks he knows about how the world works is that it is a multi-layered hierarchy in which mental features are dependent upon and determined by micro-physical properties. In explicating the implications of his supervenience naturalism for free will, Searle claims that dualism is implausible because of the problem of mind– body causal interaction. In his estimation, mind-to-body causation is deeply mysterious, if not incoherent. Kim concurs with this point. According to him, there is "the sheer impossibility of coherently imagining the details of what might have to be the case if some nonphysical agency is going to affect the course of purely physical events."[69] Acknowledging some non-physical agency

> would force us to accept a conception of the physical in which to give a causal

> account of, say, the motion of a physical particle, it is sometimes necessary to go outside the physical system and appeal to some non-physical agency and invoke some irreducible psychophysical law. Many will find this just not credible.[70]
>
> For, as many thinkers have pointed out, it simply does not seem credible that an immaterial substance, with no material characteristics and totally outside physical space, could causally influence, and be influenced by, the motions of material bodies that are strictly governed by physical law. Just try to imagine how something that isn't anywhere in physical space could alter in the slightest degree the trajectory of even a single material particle in motion. Its inability to explain the possibility of "mental causation," how mentality can make a causal difference to the world, doomed Cartesian dualism. . . .
>
> There has been a near consenus among philosophers that the concept of mind as a mental substance gives rise to too many difficulties and puzzles without compensating explanatory gains.[71]

The heart of Kim's argument is that dualism cannot be believed because it cannot *explain* how mental causation involving an immaterial substance which is not located in space is possible. One way for a dualist to respond to challenge would be to bring the soul into the same spatial system as its physical body. For example, a dualist might assert that the soul occupies a spatial point, presumably somewhere within the space occupied by its physical body,[72] and because it does it is able to causally interact with that body. It is doubtful, however, that this response by the dualist will satisfy the supervenience naturalist. Thus, Kim approvingly quotes Princess Elizabeth of Bohemia's response to Descartes that locating the soul at a spatial point will not make any clearer how an immaterial, thinking substance can causally interact with its physical body.[73] The problem for the dualist postion according to the supervenience naturalist is that no matter 'where' the soul is located, whether in space or not in space, the dualist can only say that causal interaction between it and the physical world just happens. And this inability to say anything explanatorily informative about *how* causal interaction occurs is enough to undermine any initial plausibility the dualist view has (and, therefore, any initial plausibility libertarianism has).

As I have argued elsewhere,[74] the supervenience naturalist, in spite of all of his bluster about the problem of causal interaction for dualism, is no better off than the dualist when it comes to explaining how it is that the mental relates to the physical. Thus, Kim himself has stated, "the problem of mental causation [doesn't] go away when Cartesian mental substance [goes] away."[75] To make it evident that this is the case, it is helpful to recall the supervenience naturalist's inability to say anything positive about the explanatory nature of the determinism in the supervenience relation. What this inability consists of is an inability of the supervenience naturalist to *explain* how subvenient physical properties determine supervenient mental

properties. Thus, Searle, after affirming his non-event causal interpretation of the determination in the supervenience relation, concedes that "[w]e do not know the detail of how brains cause consciousness. . . . [W]e do not know how the system of neuro-physiology/consciousness works . . . ."[76] Though there is no knowledge of how brains non-event causally determine consciousness, Searle has no hesitation to affirm that they just do. Supervenience naturalists who reject any causal interpretation of the determinism in the supervenience relation are no more helpful in elucidating the explanatory character. As noted in the quotation earlier in this section, Heil admits that such supervenience naturalists do not provide an explanation of the deterministic connection between subvenient physical properties and supervenient mental properties. Chalmers, who is himself a supervenience naturalist, concurs: "It might be objected that [the supervenience framework] does not tell us what the connection is, or how a physical configuration gives rise to experience. But the search for such a connection is misguided. . . . [B]eyond a certain point, there is no asking 'how.'"[77] Given the supervenience naturalist's own inability to explain the nature of the determination in the supervenience relation, it is at least initially puzzling why he maintains that the dualist's inability to say anything informative about how a mind and a body causally determine events to occur in each other is a decisive reason for rejecting dualism. It would initally seem that if this inability cuts against one view, it cuts equally against the other.

In an effort to be charitable to the supervenience naturalist, one might think that his inability to elucidate the explanatory character of the determinism in the supervenience relation does not cut as strongly against his view because it is reasonable to believe that it is merely a temporary state of affairs, and that with the advance of science the supervenience determinist will come to have the requisite information to provide the desired elucidation. For example, Searle states that the mystery of how consciousness is related to neurophysiology is merely a function of our present "ignorance of how the brain works."[78] This claim, however, expresses a radically mistaken belief. Some comments from Searle himself will help clarify why this is the case.

As was pointed out earlier in this section, Searle argues for supervenience naturalism by claiming that the model for understanding the relationship between the mind and the brain is found in the distinction in physics between micro- and macro-properties of systems. In physics, the global or surface features of objects are explained by the behavior of elements at the micro level. Searle's examples are solidity and liquidity. The solidity of the table on which I am writing is explained by the lattice structure of the molecules of which the table is composed. Similarly, the liquidity of water in the glass in front of me is explained by the nature of the interactions between H_2O molecules. Indeed, in the case of solidity and liquidity, it is appropriate to say that solidity and liquidity are identical with the relevant micro-structures.

> Thus, to take the example of solidity, the table in front of me is solid in the ordinary sense that it is rigid, it resists pressure, it supports books, it is not easily penetrable by most other objects such as other tables, and so on. Such is the commonsense notion of solidity. And in a scientific vein one can define solidity as whatever micro-structure causes these gross observable features. So one can then say either that solidity just is the lattice structure of the system of molecules and that solidity so defined causes, for example, resistance to touch and pressure. Or one can say that solidity consists of such high level features as rigidity and resistance to touch and pressure and that it is caused by the behaviour of elements at the micro-level.[79]

When we say that solidity just *is* the lattice structure of the system of molecules, we are claiming that causal reductionism leads to property ontological reductionism wherein one property can be shown to consist of nothing but another property.[80] Such a reduction is possible because we are able to picture the relation between solidity and the lattice structure of molecules.[81] Thus, one object's solidity vis-à-vis another can be pictured as the inability of the latter to break apart the proper parts of the former which are held together by causal forces. What is important about this model of reductionism for our present purposes is that solidity and liquidity are features of whole objects which are composed of spatially related proper parts. One would expect, then, that if the mental–physical supervenience relation is a form of causal reductionism which is elucidated by the solidity–liquidity model, the mental could be ontologically reduced to the physical because it could be pictured in terms of the spatial extension of an object and the relations of its proper parts. Searle himself, however, recognizes that the mind does not seem to be a spatial object with extension:

> We are not aware in conscious experience of . . . the dimensions of our conscious experience . . . . Although we experience objects and events as both spatially extended and of temporal duration, our consciousness itself is not experienced as spatial, though it is experienced as temporally extended.[82]

In light of the fact that the mind does not seem to be spatially extended, we are unable to picture the mind–brain relation in terms of the part–whole relation. Therefore, in the case of the mind–brain supervenience relation, causal reductionism does not lead to ontological reduction. It is important to stress, however, that the inability to picture the mind–brain relation in part–whole spatial terms is not simply a temporary difficulty which we may someday overcome. We cannot overcome it because of the logical impossibility of explaining what does not seem to be made up of spatial parts in terms of such parts. Colin McGinn summarizes this impossibility in the following

way:

> It is precisely [spatially defined properties] that seem inherently incapable of resolving the mind–brain problem: we cannot link consciousness to the brain in virtue of spatial properties of the brain. . . . [C]onsciousness defies explanation in such terms. Consciousness does not seem made up out of smaller spatial processes . . . . Our faculties bias us towards understanding matter in motion, but it is precisely this kind of understanding that is inapplicable to the mind–body problem.[83]

In light of the conceptual barriers which prevent getting a property ontological reduction of the mental in terms of the physical, we are left in Searle's terms with a causal reduction of the mental to the physical, which is really no more than saying that the explanatory nature of the determinism in the supervenience relation will forever remain an inexplicable, brute fact. As Chalmer's states, "there is an *explanatory gap* between the physical level and conscious experience. If this is right, the fact that consciousness accompanies a given physical process is *a further fact*, not explainable simply by telling the story about the physical facts. In a sense, the accompaniment must be taken as brute."[84]

In spite of acknowledging that the nature of the determinative relationship between the physical and the mental is in principle explicable on their own view, supervenience naturalists continue to maintain that dualism is not a viable philosophical position. The following are representative comments from Kim:

> You want to raise your arm, and your arm goes up. Presumably, nerve impulses reaching appropriate muscles in your arm made those muscles contract, and that's how the arm went up. And these nerve signals presumably originated in the activation of certain neurons in your brain. What caused those neurons to fire? We now have a quite detailed understanding of the process that leads to the firing of a neuron, in terms of complex electrochemical processes involving ions in the fluid inside and outside a neuron, differences in voltage across cell membranes, and so forth. All in all we seem to have a pretty good picture of the processes at this microlevel on the basis of the known laws of physics, chemistry, and biology. If the immaterial mind is going to cause a neuron to emit a signal (or prevent it from doing so), it must somehow intervene in these electrochemical processes. But how could that happen? At the very interface between the mental and the physical where direct and unmediated mind–body interaction takes place, the nonphysical mind must somehow influence the state of some molecules, perhaps by electrically charging them or nudging them this way or that way. Is this really

> conceivable? Surely the working neuroscientist does not believe that to have a complete understanding of these complex processes she needs to include in her account the workings of immaterial souls and how they influence the molecular processes involved. . . . Even if the idea of a soul's influencing the motion of a molecule . . . were coherent, the postulation of such a causal agent would seem neither necessary nor helpful in understanding why and how our limbs move.[85]

In this passage, Kim once again questions the possibility that an immaterial mind might causally influence the physical world. For the sake of discussion only, he concedes that this is possible, but then adds that the inclusion of such a causal agent is neither necessary nor helpful to a neuroscientist who is trying to understand why, for example, your limb moves. According to Kim, the neuroscientist does not need to include reference to an immaterial soul to have a complete understanding of the molecular processes (their movements) involved in raising your arm.

Contrary to what Kim's comments suggest, it is intuitively plausible to think that the neuroscientist, like anyone else (i.e. *qua* ordinary human being), does need to include reference to your reason (purpose) for raising your arm and the fact that you, a soul, caused it to move (rise) in order to have a *complete* explanation of why it moved as it did. Knowing only the neural (physical) story is not enough for such an explanation.[86] Kim is also mistaken in suggesting that the neuroscientist, *qua* scientist, is seeking a complete understanding of your arm's movement. What the neuroscientist seeks, *qua* scientist, is a knowledge of how micro-*physical* entities of various kinds relate *causally* to each other in virtue of their causal powers and liabilities. Thus, as Chalmers points out:

> Basic particles . . . are largely characterized in terms of their propensity to interact with other particles. Their mass and charge is specified, to be sure, but all that a specification of mass ultimately comes to is a propensity to be accelerated in certain ways by forces, and so on. . . . Reference to the proton is fixed as the thing that causes interactions of a certain kind, that combines in certain ways with other entities, and so on . . . .[87]

The limitation of the neuroscientist's investigation to causal relations among micro-physical entities has two important implications for dualism and free will. First, from the fact that the the neuorscientist seeks to acquire knowledge about various micro-physical entities and their forces (exercised powers) to move other micro-physical entities in various ways, it does not follow that he is required to presuppose the non-existence of souls which also have the power to move those other micro-physical entities in those ways. If a scientist makes such an assumption, he is not making it *qua*

scientist but *qua* naturalist.

Second, the propensities (liabilities) of micro-physical entities to be moved in certain ways about which the neuroscientist seeks to learn are inherently *conditional* in nature. What Chalmer's describes as a propensity of a particle to be moved is a liability of it to be moved which is such that if it is actualized by an exercised power of another entity (whether physical or non-physical), it will be necessitated to move in a certain way. There is nothing, however, in the nature of that liability of that particle which requires that it be actualized in accordance with physical deterministic laws such that the physical world is closed to causal influence by non-physical entities. Hence, the actualization of that liability by a soul on an occasion when it makes an undetermined choice for a reason is not excluded by a scientific study of that liability. And it is precisely on those occasions when arm movements occur for purposes that a neuroscientist will reasonably believe that the originative micro-physical movements are traceable to the causal activity of a soul.

IV

Naturalists such as Searle and Kim believe that libertarian freedom implies the truth of dualism. Are they correct? Like them, I am inclined to think that it does, though it has been argued that it does not. In an effort to avoid dualism and yet not be a naturalist, some have tried to develop a *via media* in the form of property emergentism. On such a view, when appropriate basal conditions obtain, higher-level, irreducible mental properties of an agent emerge which take on lives of their own. Thus, an emergentist who is also a libertarian might maintain that the power to choose is an emergent property which emerges along with the emergent causal power to produce movements in basal objects such that when an agent exercises his emergent power to choose for a purpose, he also exercises his emergent causal power to produce events in or involving the relevant basal objects. O'Connor has described such an emergentist view in terms of supervenience. On his account, an emergent property *P* is an emergent causal property (power) of a mereologically complex object *O* iff:

(1) *P* supervenes on properties of the parts of *O*;
(2) *P* is not had by any of the object's parts;
(3) *P* is distinct from any structural property of *O* (A property, *Q*, is structural iff proper parts of particulars having *Q* have some property or properties not identical with *Q*, and this state of affairs is, in part at least, constitutive of the state of affairs of the particular's having Q. In terms of Searle's ontological reducibility, solidity and liquidity are structural properties.); and
(4) *P* has direct ('downward') determinative influence on the pattern of behavior

involving *O*'s parts.[88]

As O'Connor notes, there are ways to avoid acknowledging the existence of emergent properties. For example, one might posit the presence of additional and hitherto undetected micro-properties which manifest themselves in highly complex systems of certain kinds by giving rise to *P* in such a way that *P* is constituted by them (thus, *P* is a structural and not an emergent property). O'Connor responds to this proposal by saying that "the only motivation one could have for postulating a (rather elusive) micro-property is a very strong methodological principle to the effect that one is to avoid emergentist hypotheses at all costs, which by my lights is not a reasonable one."[89]

Four points are relevant to keep in mind when evaluating the emergentist position. First, the emergentist view does not make any clearer than dualism how the mental and physical are related. For example, according to the emergentist, how a psychological property such as *P* emerges from the physical "will forever remain a mystery; we have no choice but to accept it as an unexplainable brute fact."[90] Moreover, once it has emerged, the direct downward determinative influence of *P* (see condition (4)) will be no more intelligible than the causal activity of a soul on a physical particle.

Second, in the spirit of O'Connor's emergentist response to the position that one can avoid emergentism by positing the existence of hitherto undetected micro-properties which are the structural components of the emergent property, it is plausible for the dualist to respond to the emergentist that the only motivation for postulating emergent properties is the very strong methodological principle that one is to avoid dualism at all costs.

Third, emergentism is incompatible with the conviction that teleological explanation is a deep form of explanation. That is, in so far as emergentism retains the dependency element of the supervenience thesis (see condition (1)), it is incompatible with the idea that purposes of agents are grounded in bedrock properties of bedrock entities of the world.

Finally, the emergentist view conflicts with, while the dualist position is supported by, the point made in Section III that the self seems to be simple in nature (i.e. the self seems to be an indivisible entity which is not comprised of proper parts). Our consciousness of ourselves is of entities which do not seem to be spatially extended in the sense of our having proper parts which occupy distinct regions of space. According to the emergentist, however, an object which exhibits an emergent property *P is* constituted by proper parts (see condition (2)) which, presumably, occupy distinct spatial regions . Hence, an emergentist account of free will cannot be easily reconciled with the experience of ourselves as simple entities.

V

As soft naturalists, supervenience naturalists often point out that their form of naturalism is preferable to hard naturalist views because it preserves the reality of consciousness. For example, Chalmers describes eliminativist views about consciousness "as being in conflict with the manifest facts."[91] Searle claims that hard naturalist views are false by *reductio ad absurdum* because they entail that our ordinary conscious mental life does not exist.[92] He adds that proponents of one such view, namely, functionalism, do not need refutation but help,[93] and asks "[H]ow did we get into this mess?"[94] According to his own diagnosis, the reason why hard naturalists put forth such absurd views is that they "have a terror of falling into Cartesian dualism."[95] Nevertheless, Searle and other supervenience naturalists are like hard naturalists to the extent that they are willing to deny the reality of a central element of our conscious life: choice or free will. Of them, it is appropriate to ask, 'How did they get into this mess?' As was seen in Section III, they get into this mess because they believe that free will implies dualism. Thus, someone like Searle endorses his supervenience naturalism because he, like the hard naturalists he chides, has a terror of falling into Cartesian dualism. I conclude, therefore, that in light of what naturalists themselves say it is plausible to believe that libertarian freedom is incompatible with the truth of naturalism because it implies the truth of dualism. Thus, if one is a libertarian, one cannot consistently be a naturalist.[96]

Notes

1 David Papineau, Philosophical Naturalism (Oxford: Blackwell, 1993), 1.
2 Ibid., 30f.
3 Ibid., 31, footnote 26. The emphasis is Papineau's.
4 D. M. Armstrong, "Naturalism, Materialism and First Philosophy," *Philosophia* 8 (1978): 261–76.
5 Ibid., 261.
6 Ibid., 262.
7 Eliminativists in the philosophy of mind are hard naturalists.
8 David Chalmers, *The Conscious Mind: In Search of a Fundamental Theory* (Oxford: Oxford University Press, 1996), 158–9.
9 Ibid., 124–9.
10 John Heil, *The Nature of True Minds* (Cambridge: Cambridge University Press, 1992), 4–5. The emphasis is Heil's.
11 "Naturalism is suspicious of Cartesian dualism: The mind's place is in, not adjunct to, nature." Ibid., 11. Cf. Papineau, *Philosophical Naturalism*, 1, 27; and Chalmers, *The Conscious Mind*, 124–5.
12 John Searle, *Minds, Brains and Science* (Cambridge, Mass.: Harvard University Press, 1984), Chapter Six.
13 If we let 'N' symbolize 'Naturalism', 'Psy' symbolize 'Psychological', 'Spsy' symbolize

'Supervenience of Psychological', 'D' symbolize 'Dualism', and 'C' symbolize 'Choice', the argument of Section III is as follows: (1) N→ (~Psy v Spsy) (i.e., naturalism implies either hard naturalism or soft naturalism), (2) (~Psy v Spsy)→ ~D, (3) ~D→ ~C. The libertarian asserts C and, therefore, ~N.

14 Carl Ginet, *On Action* (Cambridge: Cambridge University Press, 1990), 13f.

15 It is appropriate to point out that libertarians are not the only ones who discuss the issue of choice in terms of feeling. Non-libertarians do also. For example, Richard Double (*Metaphilosophy and Free Will* [Oxford: Oxford University Press, 1996], p. 42) asks about the reality of libertarian choice in the following terms: "How much evidential weight should we accord to our 'feelings' that we can choose otherwise than we do . . . ? . . . And how certain is it that we actually feel the way we say we do when we claim that our introspective awareness shows that we *feel* that we are able to choose otherwise . . . ?" The emphasis is Double's.

16 Wilfrid Sellars ("Philosophy and the Scientific Image of Man," in *Metaphysics*, ed. by Ronald C. Hoy and L. Nathan Oaklander [Belmont, CA: Wadsworth Publishing Company, 1991], 509) has pointed out "the mistake of supposing that in self-awareness conceptual thinking presents itself to us in a qualitative guise." Jaegwon Kim (*Philosophy of Mind* [Boulder, Colo.:Westview Press, 1996, 14]) says the following about intentional or 'contentful' states of mind: "Do these mental states have a phenomenal, qualitative aspect? We do not normally associate a specific feel with beliefs, another specific feel with desires, and so on." See also 158f. Finally, Zeno Vendler (*Res Cogitans* [Ithaca, N.Y.: Cornell University Press, 1972], 156) comments as follows: "For an asthmatic patient the feeling of his breathing is likely to be unpleasant or even painful; his opinion, on the other hand, if it is an opinion, that he is breathing, or that he is breathing painfully, is neither unpleasant nor painful. Moreover, whereas those feelings are felt in the chest and the throat of the patient, his thought that he is breathing, or that he is in pain, is not felt in any part of the body. Thoughts, in fact, are not 'felt' or 'experienced' at all."

17 Harry Frankfurt, "The Problem of Action," *American Philosophical Quarterly* 15 (1978): 157. Frankfurt makes this epistemological point with respect to bodily actions but it is equally, if not more, relevant for mental actions.

18 Myles Brand (*Intending and Acting* [Cambridge, Mass.: MIT Press, 1984], 17–23) and Irving Thalberg ("Do Our Intentions Cause Our Intentional Actions?," *American Philosophical Quarterly* 21 [1984]: 249–60) are examples of causal theorists who do this.

19 Double, *Metaphilosophy and Free Will*, 58.

20 Ibid.

21 "It is the *contents* of beliefs and desires (and not that she has the beliefs and desires) which serve as an agent's reasons for acting, while it is by being the content of *beliefs* and *desires* that these contents manage to serve as an agent's reasons." Geoffrey Sayre-McCord, "Explanations and Reasons," *Philosophical Perspectives* 3 (1989): 140. The emphases are Sayre-McCord's. He does not, however, endorse an indeterministic, non-causal account of choice.

22 Thomas Nagel, a non-libertarian, says the following about a free action: "A free action should not be determined by antecedent conditions, and should be fully explained only intentionally, in terms of justifying reasons and purposes" (*The View from Nowhere* [Oxford: Oxford University Press, 1986], 115). Robert Kane, who is a libertarian, goes further and maintains that "many libertarians (and their critics) have [actually] held that incompatibilist free choices or actions cannot be given causal explanations of any kinds. . . ." (*The Significance of Free Will* [Oxford: Oxford University Press, 1996], 174). Neither Nagel nor Kane gives the names of any such libertarians, and I think for good reason: most libertarians have not defended a non-causal understanding of free will. Instead, they have believed that it is necessary to include causation in an adequate account of free will because (I think) they have

wrongly been persuaded that a non-causal account is flawed. Kane, himself, is such a causal libertarian.

23 Richard Taylor, *Metaphysics*, third edition (Englewood Cliffs, NJ: Prentice-Hall, 1983), 45.

24 Ibid., 45.

25 Ibid., 48. Taylor defends the reality of teleological explanation and its irreducibility to causal explanation in his *Action and Purpose* (New Jersey: Humanities Press, 1980), especially Chapter 10.

26 Ibid., 116.

27 Ibid., 107–11.

28 Ibid., 174.

29 Ibid., 136, 139, 194.

30 Ibid., 140, 146.

31 Ibid., 136, 139, 194.

32 I have doubts about the reality of probabilistic causation (see the third of the following responses to Kane), but for the purpose of this objection I will concede its reality.

33 Ibid., 141. The emphasis is Kane's. I say Kane's view on this issue "seems" to be what I have stated because nowhere, as best as I can ascertain, does he explicitly explain, as opposed to merely claim, why a non-causal view of choice is inadequate.

34 Ibid., 176–7.

35 Ibid., 136. "There is . . . a teleological element involved in choosing for reasons."

36 Randolph Clarke, "On the Possibility of Rational Free Action," *Philosophical Studies* 88 (1997): 45.

37 Ibid., 55, endnote 16.

38 Randolph Clarke, "Agent Causation and Event Causation in the Production of Free Action," *Philosophical Topics* 24 (1996): 36.

39 Donald Davidson, *Essays on Actions and Events* (Oxford: Oxford University Press, 1980), 9, 11, 32, and 87.

40 For a further discusson of this issue, see my "Libertarian Choice," *Faith and Philosophy* 14 (1997): 199ff.

41 Timothy O'Connor, "Agent Causation," in *Agents, Causes, & Events*, ed. by Timothy O'Connor (New York: Oxford University Press, 1995), 173–200.

42 Ibid., 191.

43 Ginet, *On Action*, Chapter 6.

44 Ibid., 145–6.

45 In written correspondence, Ginet has expressed his preference to talk in terms of the act of adopting an intention.

46 In further written correspondence, Ginet has conceded that the description 'We have an intention formed for a reason, period' is perhaps correct.

47 Cf. Kim, *Philosophy of Mind*, 9–13, 221ff.

48 "First of all, the phenomenon of consciousness arises on the macroscopic level. That is, it is only highly organized physical systems which exhibit mentality." Joseph Levine, "Materialism and Qualia: The Explanatory Gap," *Pacific Philosophical Quarterly* 64 (1983): 358.

49 Ibid., 10.

50 Strictly speaking, it should be noted that supervenience itself is not an asymmetrical relation, but a non-symmetrical, covariance relation which must be supplemented by a naturalistic affirmation of or commitment to the ontological priority of the physical in order to achieve the desired asymmetry. See Kim, Ibid., 10–11; Stewart Goetz, "Dualism, Causation, and Supervenience," *Faith and Philosophy* 11 (1994): 92–108; and Heil, *The Nature of True Minds*, 98–102. For the sake of discussion, however, I treat supervenience as if it were itself an

asymmetrical relation.

51 Kim, *Philosophy of Mind*, 149.

52 "There is a different sort of determinism associated with naturalism that, while very much alive today, has not found its way into discussions of human freedom. It is a sort of determinism according to which the microphysical world determines the distribution of the higher level properties of material beings, adumbrated in various popular supervenience theses," J. A. Cover and John O'Leary-Hawthorne, "Free Agency and Materialism," in *Faith, Freedom, and Rationality*, ed. by Jeff Jordan and Daniel Howard-Snyder (Lanham, Md.: Rowman & Littlefield, 1996), 47.

53 A hard naturalist who claims that libertarianism is an invention of philosophers is Daniel Dennett. See his *Elbow Room: The Varieties of Free Will Worth Wanting* (Cambridge, Mass.: MIT Press, 1984). His approach to dealing with difficult metaphysical issues in the philosophy of mind is just to deny that any reasonable person believes what creates the problem, which in this case is libertarian free will.

54 Searle, *Minds, Brains and Science*, 87–8.

55 Ibid., 93.

56 John Searle, *The Rediscovery of the Mind* (Cambridge, Mass.: MIT Press, 1992), 124–5.

57 Searle, *Minds, Brains and Science*, 92. The emphases are mine. Searle recognizes the implications of supervenience naturalism for free will while others do not notice it or ignore it. For example, Kim (*Philosophy of Mind*, p. 12.) states that "very few of us would think that there can be mental events . . . that float free . . . of physical processes." Yet, he is aware that supervenience naturalism seems to imply that the mental is epiphenomenal and that "Few philosophers have been self-professed epiphenomenalists" (ibid., 129). But why are so many philosophers reluctant to endorse epiphenomenalism? The most plausible answer is precisely that at least on some occasions (i.e., when we choose) the mental does seem to float free from the physical and determines what happens in it.

58 John Searle, "The Mystery of Consciousness," *The New York Review of Books*, November 2, 1995, 60: "Look around you at the objects in your vicinity and think of the causal explanation of the fact that the table exerts pressure on the rug. This is explained by the force of gravity, but gravity is not an event. Or think of the solidity of the table. It is explained causally by the behavior of the molecules of which the table is composed. But the solidity of the table is not an extra event, it is just a feature of the table. Such examples of non-event causation give us appropriate models for understanding the relation between my present state of consciousness and the underlying neurobiological processes that cause it." Contrary to what Searle asserts, (the force of) gravity is a continuous event (or a continuous series of events). Moreover, as I will have need to point out again in a moment, not only is the solidity of the table not an extra event, but also it is not an extra feature of the table. It is *identical* with the structure of the molecules which comprise the table.

59 Kim, *Philosophy of Mind*, 149–50. See also his "Mental Causation in Searle's Biological Naturalism," *Philosophy and Phenomenological Research* 55 (1995): 194.

60 Cf. Heil, *The Nature of True Mind*, 102. The emphases are Heil's.

61 Thus, while classical naturalism was incompatible with libertarianism because of its deterministic nature, it was more fundamentally incompatible with libertarianism because of its exclusion of teleological explanation. J. P. Moreland recognizes naturalism's exclusion of teleology in his forthcoming "Naturalism and Libertarian Agency," *Philosophy and Theology* 10 (1997): 351–81.

62 Searle, *The Rediscovery of the Mind*, 13.

63 Kim, *Philosophy of Mind*, 232–3. See also "The Myth of Nonreductive Materialism," *Proceedings and Addresses of the American Philosophical Association* 63 (1989):47: "Is the abandonment of the

causal closure of the physical domain an option for the materialist? I think not: to reject the closure principle is to embrace irreducible nonphysical causes of physical phenomena. It would be a retrogression to Cartesian interactionist dualism, something that is definitive of the *denial* of materialism." The emphasis is Kim's.

64 *The Philosophical Works of Descartes*, vol. I, trans. and ed. by Elizabeth S. Haldane and G. R. T. Ross (Cambridge: Cambridge University Press, 1967), 196.

65 Nagel, *The View from Nowhere*, 43. Cf. Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1987), Part Three.

66 David Braine, *The Human Person* (Notre Dame: University of Notre Dame Press, 1992), 315.

67 Searle, *The Rediscovery of the Mind*, 91.

68 Ibid.

69 Jaegwon Kim, "Epiphenomenal and Supervenient Causation," in *Midwest Studies in Philosophy*, volume IX, edited by Peter A. French, Theodore E. Uehling, Jr. and Howard Wettstein (Minneapolis: University of Minnesota Press, 1984), 266.

70 Ibid.

71 Kim, *Philosophy of Mind*, 4.

72 For a discussion of this alternative, see Philip L. Quinn's, "Tiny Selves: Chisholm on the Simplicity of the Soul," in *The Philosophy of Roderick M. Chisholm*, ed. Lewis Edwin Hahn (La Salle, Il: Opern Court, 1997), 55–71.

73 Kim, *Philosophy of Mind*, 131.

74 Goetz, "Dualism, Causation, and Supervenience," 99–105.

75 Kim, *Philosophy of Mind*, 4.

76 Searle, *The Rediscovery of the Mind*, 89–102.

77 Chalmers, *The Conscious Mind*, 170.

78 Searle, *The Rediscovery of the Mind*, 102.

79 Searle, *Minds, Brains and Science*, 21–2.

80 Searle, *The Rediscovery of the Mind*, 113, 115.

81 Ibid., 102.

82 Ibid., 105, 127. Levine ("Materialism and Qualia: The Explanatory Gap") argues that the mind/body problem is resistant to explanation in terms of magnitudes and mechanisms. While 'Heat is molecular motion' is intelligible because heat can be explained as motions of molecular extended parts of an object which can causally produce motions in parts of other objects, 'Pain is the firing of C-fibers' is unintelligible because pain is a simple *quale* which cannot be explained in terms of an organized system of parts.

83 Colin McGinn, *The Problem of Consciousness* (Oxford: Basil Blackwell, 1991), 11–12, and 18, footnote 21. Chalmer's points out the irrelevance of Searle's solidity–liquidity model for understanding how consciousness is related to the physical: "It seems clear that this is a false analogy, however. Given all the microphysical facts about a particular batch of H2O, it is logically impossible that those facts could hold without liquidity being instantiated. . . . The mircrophysical features do not *cause* liquidity; they *constitute* it. This is entirely different from what is going on in the case of consciousness, so the analogy fails. Consciousness is ontologically novel in a much more significant way than liquidity" (*The Conscious Mind*, 130). The emphases are Chalmer's.

Galen Strawson (*Mental Reality* [Cambridge, Mass.: MIT Press, 1994]) recognizes the bruteness or "That's just how things are" (pp. 83–4) nature of the mind–brain or experience–non-experience relation, and says that he would "ban the word 'supervenience.' It has seemed to promise much, . . . but it has been of no real use at all," "Replies to Noam Chomsky, Pierre Jacob, Michael Smith, and Paul Snowdon," *Philosophy and Phenomenological Research* 58 (1998): 481, footnote 49. He espouses, however, a materialist's faith that some future physics

about which we are presently ignorant will provide an account or theoretical understanding of how experience (what-it's-likeness) is realized in the brain (*Mental Reality*, 81 ; 84–93). Given the simplicity (lack of part–whole structure) of experience, it is impossible to understand how any future physics will be able to do what Strawson hopes it will do. Why not accept dualism? Because it "faces well-known problems" (ibid., 43). Not surprisingly, Strawson denies the reality of libertarian free will. See his *Freedom and Belief* (Oxford: Clarendon Press, 1986).

84 Chalmers, *The Conscious Mind*, 107. The emphases are Chalmer's.

85 Kim, *Philosophy of Mind*, 131–2.

86 One is reminded, here, of the famous passage from Plato's *Phaedo* 98c-99a, where Socrates criticizes Anaxagoras for claiming that mind causes all things and then adducing only physical causes: "It seemed to me that he was just about as inconsistent as if someone were to say, The cause of everything that Socrates does is mind – and then, in trying to account for my several actions, said first that the reason why I am lying here now is that my body is composed of bones and sinews, and that the bones are rigid and separated at the joints, but the sinews are capable of contraction and relaxation, and form an envelope for the bones with the help of the flesh and skin, the latter holding all together, and since the bones move freely in their joints the sinews by relaxing and contracting enable me somehow to bend my limbs, and that is the cause of my sitting here in a bent position. Or again, if he tried to account in the same way for my conversing with you, adducing causes such as sound and air and hearing and a thousand others, and never troubled to mention the real reasons, which are that since Athens has thought it better to condemn me, therefore I for my part have thought it better to sit here, and more right to stay and submit to whatever penalty she orders. Because, by dog, I fancy that these sinews and bones would have been in the neighborhood of Megara or Boeotia long ago – impelled by a conviction of what is best! – if I did not think that it was more right and honorable to submit to whatever penalty my country orders rather than take to my heels and run away."

87 Chalmers, *The Conscious Mind*, 153.

88 Timothy O'Connor, "Emergent Properties," *American Philosophical Quarterly* 31 (1994): 93–6.

89 Ibid., 99.

90 Kim, *Philosophy of Mind*, 229.

91 Chalmers, *The Conscious Mind*, 164.

92 Searle, *The Rediscovery of the Mind*, 8.

93 Ibid., 9.

94 John Searle, "Consciousness, Explanatory Inversion, and Cognitive Science," *Behavioral and Brain Sciences* 13 (1990): 585.

95 Ibid., 13.

96 I want to thank J.P. Moreland for reading an earlier draft of this paper and making helpful comments.

Part 3

Value theory

8 Naturalism and morality

John E. Hare

The moral gap

The project of this essay is to point to a certain kind of incoherence in the attempt to reconcile morality with naturalism, and then to discuss what we need if we are going to avoid this incoherence.[1] I will argue that morality, as it has been conceived in modern Western philosophy, has a three-part structure, which I will call the structure of the moral gap.[2] I will focus on Kant, because I think he is the greatest and most influential moral philosopher of the modern period. I will not be talking about his work outside moral philosophy. In the Kantian structure, there is, first, the moral demand. I will stress the demand for impartiality, which is the demand to give oneself the same weight in moral thinking that one gives to any other human being. There are, second, our natural capacities which are unequal to the moral demand. They are unequal because they give us the tendency to partiality, the tendency to give ourselves more weight than morality allows. For example, they give us the tendency to desire power and prestige, both of which require (if we are to have them) that others do not have them or have them less than we do. The third part of this structure is a possible holy being, whose functioning is not limited in the way ours is, and who is seen as the source of the moral demand. Thus Kant says we should recognize our duties as God's commands.[3]

This structure of the moral gap presents a certain problem, one which Kant saw in terms of "ought" and "can." If it is the case that we ought to do something, it must be the case that we can do it. If, therefore, morality is too hard for us, it is not something we ought to be practicing. But Kant is quite certain that we are under the moral demand. He therefore makes use of the Western theological tradition, which relates the three parts of the moral gap structure this way: it says that the third part of the structure (the possible holy being) intervenes to change the second part of the structure (our natural capacities) so that they become adequate to the first part of the structure (the moral demand).

The incoherence in the attempt to reconcile morality with naturalism is the following. If contemporary moral philosophers who have embraced naturalism want

to keep the structure of the moral gap, they will have to find a substitute for Kant's (and the tradition's) appeal to divine assistance in dealing with the problem about "ought" and "can." I shall argue that they have not found a substitute, and therefore that they are left with the incoherence of saying that we are under a demand which we are unable to meet. The essay will have three sections. In the first, I describe in more detail what the moral gap picture is like, taking it one part at a time. In the second, I discuss three strategies to be found in contemporary moral philosophy for dealing with the moral gap without God. Finally, in the third section, I return to the tradition's solution to the problem, and discuss what resources it has for making possible what Kant calls "moral faith."

The moral demand

What I want to do first is to go back to the structure of the moral gap, and describe its three parts more slowly. The first part is the moral demand, in particular the demand of impartiality. Suppose we label as the 'maxim' of an action the agent's prescription of that action together with her actual reason for the action (as opposed to the reason she *gives* for it). Then we can express the moral demand in two ways. First, the maxim of an action, if it is to be morally permissible, must be able to be willed by the agent as a universal law.[4] The connection with impartiality is as follows. The agent must be willing to eliminate from the maxim of her action any singular term (most importantly, any reference to herself). She has to be willing to prescribe that the same thing should be done by any agent to any recipient in any situation like the present one in the morally relevant respects. For example, she has to be willing that the same thing should be done to her if the present roles in the situation were reversed. The moral demand is, in this form, a version of the golden rule; but it has the additional feature of connecting with the nature of reason (which always looks for universal law, whether in science or in morality). The second way of expressing the demand requires that the agent must treat humanity, whether in her own person or the person of any other, always at the same time as an end and never merely as a means.[5] This means that the agent has to share as far as possible the purposes of all those affected by her action, and the extent of the possibility here is the extent of the moral law. That is, she has to share all the morally permissible purposes of those affected by her action, and this means she has to make them her purposes. She may not treat a waiter, for example, merely as a means to get her food from the kitchen to her table. Again, the connection with impartiality is that she has to be impartial between all the parties involved including herself. She must not ignore her own desires, but by the same token she must give equal consideration to the morally permissible desires of all those affected (and those purposes will then become her own). If she does this, she will be treating those people as having the same moral dignity as she has herself.

There is a possibility here that I am only going to mention and not discuss. This is

the possibility that the two expressions of the moral demand which I have described might diverge, under one very natural interpretation of the second. This divergence does not occur in Kant himself, since on his view it is other people only as universal-law-givers who deserve treatment with equal dignity. But if we relax this requirement, we might suppose that we could treat another person as an end in herself, even if the relation between us was grounded not in our common rationality but in, for example, blood. A mother could make her son's morally permissible purposes her own, without extending this generosity to anyone else. Would she then be treating her son as an end in himself, in line with the second formula? Not in Kant's view, and not if we want to require the eliminability of singular reference.[6] But this is a large question, which I will not need to go into further. I think a large part of morality does require this eliminability; and if there is some part of morality that does not require it, then there will be a tension within morality which morality itself may not be able to resolve.

The reading of the moral demand I have given is not peculiar to Kant. It lies at the foundation of the moral systems of his successors, though these systems differ in various ways. Impartiality is expressed by the utilitarians as the requirement that each person count as one, and no person as more than one. Utilitarianism is often seen as a competitor to Kant, but Mill himself is clear that the formula of universal law is unexceptionable, indeed admirable, from a utilitarian point of view. It is only when Kant goes on to deduce actual duties of morality that Mill finds himself in disagreement.[7] This feature of the moral demand can be found in R. M. Hare's account of universalizability, in Rawls's account of the veil of ignorance, and in Habermas's account of the autonomous will.[8] These various systems differ from each other in manifold ways. But in *this* way they agree.

Our natural capacities

The second part of the structure of the moral gap is our natural capacities, which are not adequate to the moral demand. Kant puts this point in terms of our original predisposition to good, which is overlaid by a propensity to evil.[9] Here he recapitulates the traditional structure of Christian doctrine, which teaches an original creation of good, succeeded by a fall into evil. The propensity to evil, Kant says, is both innate and imputable to us, just as original sin is both of these in Christian doctrine. It is not the desire for our own happiness that Kant is condemning here. He thinks that creatures of need, like ourselves, are supposed to want to be happy. What is wrong is the ordering of happiness above duty. This is what he calls "the evil maxim," which he contrasts with "the good maxim," which reverses this priority, and tells the agent only to do what is conducive to his happiness if this is consistent with his duty. Sometimes Kant is made to sound unduly rigorous on this point, as though he thought there was something wrong with wanting to be happy. But it is the elevation of the "dear self" and its happiness above duty that he finally condemns, although his treatment of this in

the *Groundwork* is not completely clear. We are born, Kant thinks, under the evil maxim, and this radical evil is both innate and imputable to us.

On this matter, Kant's successors have not always followed him. The nineteenth-century utilitarians, in particular, shared the general belief of their contemporaries in humanity's ability to improve itself.[10] To give one example, a historian in 1908 says, "Indeed I think the most remarkable thing about our civilization and the thought of the world to-day, is the increase in the spirit of brotherliness and of unity."[11] John Stuart Mill puts his faith in education and what he calls "opinion." He says that the principle of utility prescribes

> that education and opinion, which have so vast a power over human character, should so use that power as to establish in the mind of every individual an indissoluble association between his own happiness and the good of the whole; . . . So that not only he may be unable to conceive the possibility of happiness to himself, consistently with conduct opposed to the general good, but also that a direct impulse to promote the general good may be in every individual one of the habitual motives of action, and the sentiments connected therewith may fill a large and prominent place in every human being's sentient existence.[12]

This optimism did not survive the carnage of the twentieth century. It is hard to see that education and opinion have had the effect Mill was hoping for.

Some of the more recent utilitarians have been less sanguine, and I will take R. M. Hare again as an example. The highest level of moral thinking, in his system, is what he calls "critical thinking," and it involves knowing the preferences of all the parties affected by a decision and being completely impartial between them. He ascribes this kind of thinking to an imagined archangel, and distinguishes it from "intuitive thinking" which he ascribes to an imagined prole. He then says:

> The critical thinking of a perfect moral thinker manifests this full representation and identification. Archangels can do it, and of course God. We human beings are not gifted with so much sensitivity or sympathy, and for that reason have to do for the most part with intuitive thinking; but we have to try, if we are to do the critical thinking which would validate our intuitions.[13]

This means that critical thinking is *both* something we ought to do, in order to validate our intuitions, *and* something which we do not have the sensitivity and sympathy to do as we should. This is the structure of the moral gap.

I will return shortly to the suggestions that we can deal with the moral gap by exaggerating our capacities to meet the moral demand, or by relaxing the demand to

fit our capacities. At the moment my point is just that the picture of the moral gap fits the experience of the moral life. Neither of these two suggestions will satisfy us in the end.

The possible holy being

The third part of the moral gap structure is the possible holy being. In Kant, as I have said, this being is God. It is striking, however, that even moral philosophers who have not retained belief in a supernatural God, have retained the formal place of this possible holy being in their moral theories. This can be done either by addition to our human capacities, as in R. M. Hare's archangel, or by subtraction, as in Rawls's veil of ignorance. These two theories differ in many ways, but they are alike in this: the authoritative moral prescription is given by someone who is not in the usual human situation. This is something that requires explanation. Why should not morality be seen as a paradigmaticaly *human* institution? Nietzsche already had this thought, and made the prediction that guilt (which arises from the recognition of the gap) would decline along with the Christian God. He said, "there is no small probability that with the irresistible decline of faith in the Christian God there is now also a considerable decline in mankind's feeling of guilt."[14] It is true that some contemporary moral philosophers have conceded Nietzsche's connection between the gap picture and traditional religion, and have accordingly urged us to reject the gap picture.[15] But if this is the price for rejecting traditional religion, I think the price is too high. If the moral gap picture is right, we have to make the following choice: either we must accept the traditional religious picture of divine assistance or we must find a successful substitute for it.[16]

I do not mean to imply by anything I have said that the picture of the moral gap is unique to Christianity. It can be found in the other great monotheistic religions. It can also be found in Aristotle. He tells us that the best life would be superior to the human level, but that we ought not to follow the proverb-writers, and "think human, since you are human," or "think mortal, since you are mortal"; rather, as far as we can, we ought to be immortal.[17] I believe it can also be found outside Western thought. I am not an expert in Chinese philosophy, but there seems to be a similar picture to be found there. When I was recently teaching in China, I was referred to the thought of Chu Hsi. He uses the picture of a pearl (our natural goodness) in muddy water, and seems to have held that this nature "as it is found actually inherent in the physical endowment of an individual, . . . always strives for the ideal, but always falls short of it and cannot attain it."[18] Kant might ask the question, "If our predisposition to goodness (the pearl) is obscured by distortive desires and thoughts (the muddy water), how can we see and will clearly enough to correct them?" If this is not a picture of the moral gap, it certainly suggests this picture at first reading.

Why should we accept the principle that "ought" implies "can"? It would be one

way to disarm the problem of the moral gap to deny this principle. But the principle seems to be embedded in much of our moral practice. In the education of children, we recognize that we should not hold them accountable to standards which they are unable to meet. A child should not be blamed for knocking over her milk, if she has not yet developed the motor skills for the fine control of her limbs. It is one of the hardest tasks for teachers to make what they demand of their students fit their students' capacities. But though this is hard, we think the teachers have failed seriously if they do not do this.

It might be objected that Christian doctrine is guilty of just such a failure. Augustine says,"God bids us do what we cannot, that we may know what we ought to seek from him."[19] Luther uses the analogy of a parent teaching his child to walk. The parent tells the child to walk to him, and the child tries and falls over. Why does the parent ask? Because he wants the child to reach out for the parent's hands which are held out to help him.[20] In the same way, Luther says, God gives us his law in order that we might know that we are unable to live in a way that pleases God unless we call on God for assistance. The point is that the doctrine here does not contradict the principle that "ought" implies "can." This is because the principle does not state that "ought" implies "can by our own devices." What we are capable of (our capacities) depends on what assistance is available.[21]

Three secular strategies

The next section of this essay is a brief description of three ways that can be found in contemporary moral philosophy for dealing with the moral gap without invoking divine assistance. I do not think that these strategies work.[22] My purpose here is merely to give an indication in outline of what the alternatives are to the view I am presenting.

Puffing up the capacity

The first strategy is to deny the gap by pretending that our natural capacities are after all equal to the moral demand. I call this strategy "puffing up the capacity." I have already mentioned one variant of this strategy, which can be found in nineteenth-century utilitarianism. For Mill, there is a defect in our education and in "opinion" that prevents us living up to the utilitarian principle, and these defects can be remedied. I want now to describe some contemporary utilitarians who make much the same point. Shelly Kagan, in *The Limits of Morality*, argues that if all our beliefs were fully vivid, including especially our beliefs about the interests of others, we would tend to conform to the impartial standard that utilitarian morality requires.[23] Suppose, for example, that I am purchasing a ticket for a movie which I know in advance will not be of high aesthetic or moral character. It may occur to me that the money I am about to

spend could keep someone currently starving alive for about a week. In utilitarian (and also in Kantian terms), I have to ask whether I could continue to will the spending of this money this way if I did not know whether I was going to be the movie-consumer, or the man dying of starvation. There are many complications in considering this sort of case, and I have written about them elsewhere.[24] Utilitarianism can justify some non-essential spending on myself even in the face of this thought experiment. But most people in the affluent parts of the world in fact habitually make expenditures which are unjustified according to the utilitarian principle. Kagan's claim, then, is that if our beliefs about those who are starving were fully vivid, we would tend not to make those expenditures. I want to claim that he is wrong about this, and that his mistake arises from failing to take seriously what Kant calls "radical evil."

Kagan's argument is that we can draw on the analogy of prudence, or enlightened self-interest. Suppose I hate going to the dentist. The thought of the drill in my mouth this afternoon may weigh more with me than the thought of the long-term future health of my teeth, and I may decide to cancel the appointment. But I can be won over by the thought that if I were being prudent I would prefer the long-term health of my teeth. Abscesses are a lot worse than the drill. I can, by attending to this thought, increase my tendency to go to the dentist, and thus do what prudence dictates. So we can *both* concede that we are by nature biased in favor of the present *and* say that prudence is authoritative for us. For despite the bias, we can be motivated by the thought of what we would be motivated towards if we were prudent. We can make the counterfactual claim that if all our beliefs were vivid, including especially our beliefs about our long-term future interests, we would tend to do what prudence requires. Kagan then draws the analogy. It is true, we can concede, that we are biased towards our own interests over those of others. But we can be motivated by the thought of what we would be motivated by if our beliefs about the interests of others were as vivid as our beliefs about our own interests. And he makes the corresponding counterfactual claim (as he needs to do). If all our beliefs were vivid, including especially our beliefs about the interests of others, we would tend to conform to the impartial standpoint (to what the utilitarian principle requires). Evidence for this counterfactual claim is that when we do succeed in making the interests of others vivid, it does increase our tendency to give weight to them. Some relief organizations capitalize on this fact by linking donors with particular children in recipient countries, and send photographs and periodic reports on their progress. It is easier to give sacrificially when we are linked in this way.

Unfortunately the analogy between morality and prudence breaks down at the key point. Suppose, to modify one of Judith Jarvis Thomson's examples, there are two brothers who both love chocolate, and one of them receives a chocolate bar as a gift.[25] This example has the advantage that the chocolate-owner knows his brother well, and represents to himself quite vividly his brother's interest in the chocolate. Suppose morality requires that he share it (but we leave open, because it is not relevant for

present purposes, the question of whether *justice* in some strict sense requires it). Suppose that the chocolate-owner knows that if he were committed to doing what is morally right, he would share it. Now we have to ask whether this knowledge gives him the tendency to share. This is not merely the question whether his tendency to share is increased, but whether that tendency becomes greater than half. I think the answer is that it all depends on whether or not he is already committed to morality. Some people are and some people are not, and some people are committed in some arenas and not in others. There are again many complexities here which I will not go into. My point is just that we can know perfectly well the harm we are doing to other people, and know that if we were committed to morality we would not do it, and then go ahead and do it anyway. Consider the bullying on the school playground at recess. Perhaps, the optimist may argue, it is only *children* who behave this way. But then we need to consider the refined torture which married couples can put each other through. Knowing how much they are hurting each other can *increase* the tendency to do it. To return to the case of famine relief, I think that people can know that if they were to represent to themselves vividly the suffering of those who are starving, their tendency to give would be increased, and so they turn off the television when the pictures of starving children come on. This is because the claims of morality have, so to speak, already lost the battle.

Other recent utilitarians have adopted the approach of what is sometimes called the "objective list." Thus James Griffin, in *Well-Being*, produces a list of the things that together make human life good.[26] He aims here at what he calls a "common profile," and he calls the desire for these things "the natural structure of desire." 'Virtually all persons, when informed," he says, "want to live autonomously, to have deep personal relations, to accomplish something with their lives, to enjoy themselves." His list is completed by a fifth item, understanding, or knowing about oneself and one's world. He allows that individual differences between people can affect the degree to which and the manner in which a particular person can realize one of other of these values, but he claims that any normal person will conform to the common profile in spite of such differences. On Griffin's moral theory, we can therefore decide what to do on the basis of maximizing the satisfaction of peoples' informed desires (where "informed" desires are those that conform to the natural structure of desire).

The first thing I want to note about this theory is that there are similar lists given by non-utilitarian theorists, for example by contemporary Thomists such as Germain Grisez, Joseph Boyle and John Finnis.[27] Grisez, for example, lists the various kinds of human good as self-evident objects of pursuit: self-integration, authenticity, justice and friendship, holiness, life, knowledge, and exercises of skill. These Thomist lists differ in various ways from Griffin's, but I want to point to two. They include the relation to God as a basic human good, and they include communal goods and not merely deep personal relations. In both these ways, they reflect the Aristotelian and the Christian tradition. Griffin's list is that of a secular individualist Westerner, and this

casts doubt on the claim that he has really captured the common human profile.

More important for present purposes, however, is that Griffin and the contemporary Thomists both miss out from their lists items such as power and prestige, which ought to belong there if the claim is to capture what ordinary people in fact aim at.[28] These lists are too benign. The moral gap picture I have been presenting is one in which after the Fall we naturally aim at competitive goods, which are goods that we can only have if others do not have them or have less of them than we do. Kant says that human beings, as soon as they enter into each other's company, are besieged by envy, the lust for power, greed, and the malignant inclinations bound up with these.[29] The desire to control and dominate others may be at least as strong as the desire to understand oneself and one's world. In fact the first desire often lies behind the motive power of the second. We want the knowledge in order to produce the means of control. Even the good we aim at is distorted by the evil.

Reducing the demand

The second main secular strategy for dealing with the moral gap without invoking divine assistance is the opposite of the first. I will deal with it more briefly. The first strategy aimed to bridge the gap by exaggerating our capacities so that they fit the moral demand. The second strategy is to diminish the demand so that it fits our natural capacities. The main example I shall give of such a strategy is from the feminist ethical theory which takes its origins from Carol Gilligan's work on the ethic of care. Gilligan herself does not so much reduce the demand as add to it. But Nel Noddings, in *Caring*, writes as follows, "I am not obliged to care for starving children in Africa, because there is no way for this caring to be completed in the other unless I abandon the caring to which I am obligated."[30] Caring is, for her, a rich relation, calling for "a displacement of interest from my own reality to the reality of the other," and this is only possible for a moral agent within a comparatively small group of other people. I will call this view "extreme particularism," because it holds that our obligations are confined within the special relations to those we can care for, in her sense. There are other versions of particularism, one of which I argue for myself, which hold that our obligations include, but are not confined to, those with whom we have special relations.[31]

If extreme particularism were right, this would indeed reduce the moral demand. This is not because caring is easy. On Noddings's account, the demand to care is a very high demand. But at least we are relieved of the obligation to those starving children in Africa (though she does say that we should maintain an internal state of readiness to try to care for whoever crosses our path). The special relations are more natural sources of self-sacrifice. I do not think, however, that extreme particularism can be right, and this for three reasons. First, it does not fit our ordinary usage of the moral terms or our ordinary moral practice. When famine or floods hit, there are many

thousands of people in the richer nations who hear about it or see it on television, and give generously to people with whom they have no antecedent relations at all, and without even knowing who the recipients of their gifts are going to be.[32] Second (though this in a way begs the question), if we were to adopt the extreme particularist view, the suffering in the world would increase. Unfortunately the truly desperate in the world cannot depend for their survival or nurture on those who have special relations to them. I say this begs the question because it assumes that it is a morally bad thing that the suffering of the world increases. It would be possible for an extreme particularist to deny this, on the grounds that what is morally bad is confined to what we have an obligation to do something about, and we have obligations only to those for whom we can care. Third, even the special relations like friendship and family need an impartial sense of justice if they are to function well. The people within these relations need to look at each other not merely as objects of special favor, but also as human beings deserving no more and no less than any other human beings. There is, so to speak, a base line of respect (the recognition of what Kant calls "dignity") which special relations are often tempted to erase.[33] Though it is a rich subject, I am not going to say anything more here about the relation between universal morality, which conforms to the first expression of the moral demand I started with, and the morality internal to special relations.

Finding a secular substitute

The third main strategy for dealing with the moral gap without invoking divine assistance is to leave the moral demand and our natural capacities where the moral gap picture says they are, and then find a substitute for divine assistance to bridge the gap. There are various proposals in contemporary moral philosophy which can be understood in terms of this strategy. I will discuss one, the appeal to evolution. But there are at least two others which a fuller discussion would cover. There is the appeal to rational choice theory (where this is seen as a kind of machinery operating on our choices, like an invisible hand) and the appeal to the forces of history (as in Marxist versions of Hegel).

It may seem strange to describe evolutionary ethics as looking for a substitute for divine assistance, since the workings of genetic mutation and natural selection would seem to belong to the category of our natural capacities. This would make evolutionary ethics a proposal that our capacities are in fact equal to the moral demand. But it is helpful to see evolution as a substitute for divine assistance, because it is proposed as operating beyond our conscious control. This gives us not a *deus ex machina*, but just a *machina*. Nevertheless, it is proposed as something external to the conscious life of practical agency, and as something which saves us from the otherwise tragic and irresoluble tension between duty and inclination. In fact some proponents of evolutionary ethics have been quite explicit about the idea that evolution functions

as a replacement for God in solving the problem of the moral gap. Thus Donald Campbell proposes an account of social evolution that gives us a solution to what he sees as a biological bias that humans have towards egoism. Social evolution, he says, counteracts this bias at the level of social system preaching, which produces the demand for altruism. Campbell's idea is that this demand produces just enough countervailing pull to bring us to what he calls "the bio-social optimum," roughly halfway between the biological optimum (which is egoistic) and absolute altruism.[34] Here we have the recognition of the moral gap. Campbell himself is a naturalist, in the sense that he wants to do away with "a supernatural transcendent authority," but he also recognizes the traditional function of the third part of the moral gap picture, the holy being who assists us. He therefore proposes that those who cannot any longer believe that there is such a being do a translation exercise; they should translate talk about a supernatural God into talk about something else that can function this way, such as our ecological niche.[35] Those who can go on believing in the traditional picture should do so, he thinks, for this is socially beneficial. But this is not because, on Campbell's view, this picture is true. It is a picture produced by processes mysterious to us, but analogous to biological evolution. He urges that the fact that we do not understand the process should not prevent our using its results, just as it is rational on evolutionary grounds to continue breathing air before one understands the role of oxygen in bodily metabolism.

There are some difficulties with this view. The first is with the analogy embodied in the notion of social evolution itself. For the analogy to work for explanation there would have to be something at the social level like heterogeneous variation, systematic selection, and retention and replication of the variation. I am a mere philosopher, and do not want to claim to know that there is no such thing. (I suspect that a burden of proof point is appropriate, however. In the absence of the kind of study that shows that this kind of explanation works, we should remain skeptical about its chances.) A second difficulty is that there is a curious kind of dishonesty in the account. Campbell's kind of naturalist knows that giving the translation of theistic language in terms of a naturalistic scientific analysis will tend to diminish the effectiveness of the preachments in their original form. He can give it to carefully screened groups of his peers, but he cannot let it get out into the general public. The third difficulty is that the translation seems quite inadequate to do the job the original beliefs were supposed to do. The ecological niche is not a very plausible object for our worship, and there is a principled reason for this. In Campbell's view the niche is not merely biological. It is something altered by our actions; it is what we have as a species done with what we have been given. Our niche now includes, for example, nuclear weapons. But the object of worship, if it is to have the kind of countervailing effectiveness which Campbell wants it to have, must transcend our actions if it is to have authority over them.

One way for an evolutionary ethicist to respond to these criticisms is to drop the

notion of social evolution. I will take one example of a contemporary philosopher who does this, namely Allan Gibbard in *Wise Choices, Apt Feelings*.[36] Gibbard says that to think an act morally reprehensible is to accept norms that prescribe, for such a situation, guilt on the part of the agent and resentment (or anger) on the part of others. A moral judgment that an act is wrong is then the expression of the acceptance of such norms. We can use evolution, he suggests, but now straightforward biological evolution, to explain how it may be natural to accept the norms. We are probably biologically conditioned to anger (since it seems to be pervasive across cultures). But anger makes coordination, and hence the goods attainable only through coordination, more difficult. Since guilt leads to actions that mollify anger, such as apology and reparation, evolution gives us guilt in order to make those goods accessible.

There is a problem, however, which Gibbard recognizes, with saying simply that evolution gives us guilt. The problem is that guilt seems to be culturally specific. If an emotion is genetically determined, we can suppose that it is a physical state of the organism which Darwinian forces have selected because it stands behind a pattern of overt kinds of behavior. Perhaps anger is like this; but guilt is not. Gibbard says:

> Shame or something like it appears everywhere, and so does anger, but guilt seems culturally special. Many languages have no word for guilt, and so if the people who speak these languages do feel guilt they cannot say so. Yet if guilt and shame are distinct biological adaptations to the human condition, then they were widely present as human nature evolved, and they should appear distinctly in a wide range of present cultures.[37]

Since he is talking about the evolutionary time-scale, it is especially in the surviving hunter-gatherer societies that he would need to find guilt, and he does not. What is the difference between guilt and shame? Roughly, guilt is a response to the voluntary *doing* of wrong, and the consciousness of lack of proper motivation, whereas shame is a response to inadequacy or blemish in what one is, and the consciousness of lack of proper resources. Gibbard therefore has to say that guilt is a different kind of emotion from anger, one that is significantly a matter of peoples' own conceptions, which will vary sharply from one culture to another.

What are the cultures which produce the sense of guilt? The moral gap picture suggests that we will find guilt in those cultures which also believe in a holy being who is without human moral limitations and is seen as the source of the moral demand. This is just what we do find. This is also why Nietzsche could express the hope (in the passage already quoted) that the death of God in our culture would free us from the sense of guilt. But this connection makes a difficulty for those who want to use evolution as the ground explanation for the moral-gap kind of morality. We would

need an evolutionary account of why some cultures specialize in religious views of this kind. If we had such an account, we could add that there is a genetic programming "for broad abilities to pick up the emotional script one's culture provides." But without such an account, it does not seem that evolution can provide the required explanation.

Moral faith

In the third and final section of this essay, I want to return to the solution which Kant and the tradition offer to the problem of the moral gap. Invoking divine assistance avoids the incoherent position of putting ourselves under a demand we are unable to reach. This is the negative way to put the point. We can also say, positively, that we need to have moral faith of two kinds. We need to have the faith that we can be morally good, and we need to have the faith that being morally good is consistent with being happy. I will link the discussion of the first of these to some general remarks about Kant's relation to the traditional religion in which he was reared, and his relation to the Christian Bible.

Pure rationalism

Some interpreters have urged that Kant is a deist, in the sense that he does not accept special revelation, especially the authoritative revelation by God through the Christian Bible.[38] I think it is better to take him, in his own term, as a "pure rationalist," which is to say that he accepts special revelation but does not think that this acceptance is necessary to rational religion.[39] It is helpful to look at the prologue to the second edition of *Religion within the Limits of Reason Alone*, where he suggests that we construe revelation in terms of two concentric circles. The historical revelation (to particular people at particular times, for example through the Bible) lies in the outer circle; the revelation to reason (which Kant takes not to be historical in this sense, but to be the same in all people at all times) lies in the inner circle. People in China, for example, who have never heard of Christianity, will still receive from God, on Kant's view, this revelation to reason. Kant then has two projects. The first is to see if the contents of the outer circle are consistent with the contents of the inner. The second is to take the central items of the historical revelation, and see if they can be translated into the language of the inner circle when looked at in the light of the moral concepts. If I am right that he is a pure rationalist, we can expect him to say that the historical faith *is* consistent with the religion of pure reason, and we can expect him to demonstrate this by means of the translation project. But we can also expect him to deny that the historical faith is the only possible way to reach a saving faith. He thinks it is possible for God to decree salvation even though a person has lived outside the contingent

geographical and historical reach of the Christian scriptures. He thinks that he himself, however, and his contemporaries in Christian Europe, are amongst those for whom the Christian scriptures have been the "vehicle" of a saving faith.

Despite being a pure rationalist, however, Kant also concedes that we may have to believe some item in the outer circle, even though we cannot make use of it with our reason in either its theoretical or its practical employment. This is most importantly the case when we are dealing with what he calls "Spener's problem," referring to the great Lutheran pietist, Philipp Jakob Spener. The problem is "how we can become *other* men and not merely better men (as if we were already good but only negligent about the degree of our goodness)."[40] Kant agrees with Spener (and Luther) that human beings start out in bondage to radical evil. In Kant's terms, as I said in the first section of this paper, humans are born with an innate and imputable tendency to prefer their own happiness to their duty. Since we live, he is certain, under the moral law which makes duty prior to happiness, it must be possible for a revolution of the will to take place in us, by which the immoral preference is reversed. Otherwise we would not be able to follow the moral law. But how can such a revolution take place? If the propensity to evil is radical, it is inextirpable by human powers, "since extirpation could occur only through good maxims, and cannot take place when the ultimate subjective ground of all maxims is postulated as corrupt."[41] This is Spener's problem. And Kant's solution is also Spener's; he appeals to divine assistance. It must be God who works in us to bring about the revolution of the will. This doctrine of divine grace is something Kant thinks we have to accept if we are to sustain the moral life, but he does not think he can translate the whole of what we need into the language of the inner circle. This makes Kant's position unstable, and I will pursue the question of this tension one further step.

We see in Kant's view about the revolution of the will one kind of moral faith, the faith that "Heaven will find the means to make up our deficiency."[42] Kant is here using the Christian doctrine of justification. God sees us as just, or righteous, or well-pleasing to him, "through a purely intellectual intuition."[43] In Kant's doctrine, intellectual intuition is productive. God does not passively receive what he sees; he makes it. In Lutheran terms, justification is constitutive. As the Formula of Concord puts it, "God bestows and imputes to us the righteousness of the obedience of Christ; for the sake of that righteousness we are received by God into favour and accounted righteous." When God looks at us, he sees his Son, because he is imputing to us his Son's righteousness. Kant has a partial translation for this doctrine, by reading God the Son in this context as humanity in its moral perfection.[44] But he does not have a way to translate the claim that something which is not already human has to intervene in order to change the moral propensities with which we are born. If this is right, then either we need to supply the translation which Kant fails to give us, or we have to acknowledge that we have to go outside the inner circle in order to find the beliefs

which can sustain the moral life.

Providence

The moral agent has to believe that her capacities have been transformed *inside*, so to speak. That is the first kind of moral faith she has to have. The second is that she has to believe that the world *outside* is the kind of place in which happiness is consistent with a morally good life. There is a natural objection here, that moral duty is supposed to be a sufficient end all by itself, a "pure" end, whether it leads to happiness or not. To put this more strongly, is not there something disreputable about keeping hold of the desire to be happy in the face of the majesty and dignity of the moral demand? The second kind of moral faith seems to require the moral agent both to be committed to morality as a pure end, excluding self-interest as a motivation, *and* to care that morality is consistent with her own happiness. Kant's view, however, is not that there is anything disreputable about the desire to be happy as such. He thinks this is our fixed nature.[45] What is disreputable is subordinating our duty to our happiness. We can say that the moral agent who has moral faith foresees her happiness as consistent with living a life of duty. So far there is nothing wrong. It would only be wrong if she did her duty *in order to* be happy, or if she viewed her duty as a *means* to her happiness.

I am going to proceed by identifying within this second kind of moral faith two component beliefs, which I will call the less ambitious belief in providence and the more ambitious belief in providence. First, there is the less ambitious belief that the world could be such that every person is morally good and every person is happy. This would be the best kind of world, and faith in providence contains the belief that our world could become such a world. Second, there is the more ambitious belief that the actual world is so ordered that a person's own virtue is consistent with her own happiness, whether other people are virtuous or not. I will proceed by showing, first, that the less ambitious belief can be derived from other more familiar beliefs required by the practice of morality. Then I will try to show that the less ambitious belief is not enough, and the more ambitious belief is needed as well.

The less ambitious belief may seem odd, not the kind of belief that most people have or have even considered. What is the importance of the belief that the world could be such that every person is both morally good and happy? I think this possible state of the world functions for us as an ideal, which inspires us towards engagement with the world as it actually is. But for it to function in this way as an ideal, this state of the world does have to be possible. There is an expression of this ideal in Psalm 85:10, in the vision that righteousness and peace (in Hebrew *shalom*) will kiss each other. This idea of peace is a world in which we are all happy; but "happy" is too flimsy a word. It is not merely that in such a world we have what we want, but that what we want is what it is morally good to want. That is why it is *righteousness* and peace that kiss each other. This is a vision by the psalmist of the full kingdom of God, and it has the power to

sustain us as we try to bring the actual world closer to the vision than it now is. When we do see glimpses of the kingdom, the vision allows us to hold them together into a pattern, and to recognize them as significant.

We can derive the less ambitious belief in providence from two other beliefs that a moral agent has to have about her own life and the lives of other people. She has to believe, first, that she can, over the course of her life as a whole, will the good.[46] This is made problematic by the gap between the moral demand and our initial capacities, and I have discussed various secular solutions to this problem which I claim do not succeed. The second thing she has to believe, I think, is that she can achieve, a significant proportion of the time, the good things she aims at. She has to believe this, because otherwise there would be no point in aiming at them. To make this vivid, consider what would happen in our moral life if there were the moral equivalent of Descartes's evil genius or demon. In Descartes, the point is about the state of our knowledge. The evil genius is imagined to exercise his fiendish distortion so that all our attempts at true belief about the world end up with falsity. But suppose the evil genius concerned himself directly with our moral lives, and brought it about that whenever we tried to do good we ended up doing harm. I think we would stop trying to do good.[47] This second belief is required by the first. I do not think it would be possible, over the course of an agent's life as a whole, to will the good, if she did not believe that she could achieve that good a significant proportion of the time. The manipulations of the evil genius, if she knew about them, would stop her trying to do good. And if she could not try to do good, she could not will the good, though she could still *wish* that the good would occur independently of her will. An agent's will is formed in deliberation about what to do, and she could not will the good if she thought her moral deliberation would constantly be frustrated in this way.

These two beliefs about her own life apply also to the lives of other people. The moral agent has to believe that *they* are capable of willing the good, and that *they* would then be capable of achieving, a significant proportion of the time, the good things they would be aiming at. Why does a moral agent have to believe these things about other people? One answer to this is that morality requires us to think of each other as people who could be good people. The moral agent has, we might say, to *respect* other people. This does not mean she has to regard them as virtuous. But she cannot respect them if she does not regard them as capable of virtue. This does not mean, either, that she has to believe them capable of virtue by their own resources. To respect another person requires believing that he has enough good in him so that *given the assistance that is in fact available*, he could be a good person. It is Kant's argument that this requirement means that we have to think God's assistance is available.

I said that the moral agent has to believe that if everyone were virtuous, they would be able to accomplish, a significant proportion of the time, the good things they try to accomplish. Another reason for this is that we have to think, if we are to be able to deliberate successfully, that other people can, much of the time, carry out what they

decide to do in *their* deliberations. This is because our intentions are massively interconnected with other peoples' intentions. It is hard to think of intentions we might have that do not require the trust that others have the ability to carry out most of what they choose to do. Consider, for example, the intention which I formed when I started writing this essay, to have it published in this book. The publication of such a book requires deliberation and action by all the people who produce, market, distribute and sell books. For all the people involved in this way, consider then a second larger group, all the people relied on by the first group in their daily lives. Then consider a third larger group relied on by the second, and so on. In the end we will get to almost everybody.[48]

Now if we add together these two beliefs as they apply to herself and to others, we get the less ambitious belief in providence which I started with. The best way to see this is to suppose that these various possibilities are actualized. Suppose everyone is in fact virtuous, and everyone in fact achieves much of the time the good things aimed at. Then roughly everyone will be happy. In such a world the happiness of others is what everyone will be aiming at; and if they are achieving what they are aiming at, they will be achieving each other's happiness. Kant calls this "the idea of self-rewarding morality."[49] It is because we are collectively virtuous, on this vision of the good, that we collectively secure each other's happiness. If everyone is virtuous, there can still be tidal waves and arthritis, which happen independently of our deliberation. But such natural evil will not be, for most people, sufficient to destroy their happiness. This is because they will be embedded in loving relationships with other people, and surrounded with compassionate and competent care-givers. What we have here is an idea of *self*-rewarding morality, because the idea is that it is everyone's virtue which results in (roughly) everyone's happiness.

But the less ambitious belief in providence is not enough. Here is the conclusion of the argument. We have to be able to persevere in morality even if we do not believe that most people are in fact morally virtuous. We need what I called the *more* ambitious belief in providence, the belief that the actual world is so ordered that a person's own virtue is consistent with her own happiness, *whether other people are virtuous or not*. I have noticed a very widespread gloom about the decline of virtue, and decreasing trust (at least in the countries I know well) in the general goodwill of other people. What is important for my present purposes, however, is not how virtuous other people actually are, or even what the general belief is about the general virtue. The important thing is that the moral agent's commitment to morality does not depend upon her belief in the virtue of others.

Consider the fact that morally good agents try to teach their children to be moral, and also want their children to be happy. If they thought that being virtuous would make their children miserable, they would be more ambivalent about teaching virtue than they actually are. But they persevere in the attempt to bring up their children to be morally good people even though some of them do not think their children will be

living in a society in which most of their fellows are morally good people. What happens to moral education in periods of great social turbulence and upheaval, when other people's behavior becomes quite unpredictable? Do morally good parents go on trying to teach their children to be good people? My point is not that they do, or that they do not, but that morality *requires* that they do. Since good parenting also requires caring for their children's happiness, it requires something like the more ambitious belief in providence I have been talking about. If the idea of *self*-rewarding morality were the only kind of providence we could believe in, this requirement of perseverance would be quite mysterious. What lies behind such a requirement is surely a belief that the world is so ordered that when their children grow up, they can be both morally good and happy, and that this is secured not by general human virtue but by something else. The nature of this 'something else' is often, I think, left indeterminate. Eventually we need a more complete account of what this kind of trust in the moral order of the world requires.

The experience of evil

This is especially true because the more ambitious belief in providence is not clearly supported by experience. That is why moral *faith* is the right term here. Experience gives us all sorts of cases of morally bad people who are to all appearances happy, and morally good people who are to all appearances unhappy. Our experience is thus consistent with a much bleaker picture of how the world is. This is the kind of picture which Bernard Williams locates in the world-view of the ancient Greeks. He says, of Sophocles and Thucydides, that they represent "human beings as dealing sensibly, foolishly, sometimes catastrophically, sometimes nobly, with a world that is only partially intelligible to human agency and in itself is not necessarily well adjusted to ethical aspirations."[50] He identifies this as the ancient sense of tragedy and the vulnerability of human life to the caprice of fortune. There is here a competitor to moral faith and a much darker picture of our destiny. Actually I think Williams has got Sophocles and Thucydides wrong. In Sophocles's tragedies, Oedipus and Ajax are made to suffer by an agency which they do not understand. But it is anachronistic to think of this in Nietzschean terms. Our failures to understand what is happening to us do not license the conclusion that there is no moral order. They require a humility about our own understanding of the object of our faith, which Kant also urges upon us. Thucydides, though he is often presented as cynical, in fact had a deep belief that certain kinds of moral decay would lead to certain kinds of catastrophe. He wanted to lay out the symptoms he had observed in order to give his successors advanced warning. But these exegetical questions about the ancient Greeks are not the main point. Williams does present us with an alternative to moral faith, and that is what is important for now, rather than the question whether he is reading Nietzsche back into the ancient world, because he himself finds the Nietzschean picture congenial.

What Williams does is to tie the rejection of moral faith to the rejection of the traditional picture of the moral demand I started with, especially the demand for impartiality. My argument has been that moral faith is morally necessary; necessary, that is, if the commitment to morality is to be sustained. But someone who rejects the picture of the moral demand I have been giving can use my argument to his advantage by reversing it, or turning it on its head. Williams, following Nietzsche again, wants to discard the Kantian picture of the moral demand. He belongs, therefore, under the second strategy I discussed in the second section of this essay, the strategy of reducing the demand. He also thinks that the evidence of the lack of fit between virtue and happiness in this life is overwhelming. It is so strong that only a fool or a bigot would go on believing that the world is so ordered that we can both meet the moral demand in its traditional form and be happy. Moral faith is not possible in the kind of life that Williams finds admirable. Rather, it is the product of weakness or self-deception. But if the traditional picture of morality requires moral faith, and if moral faith is not possible in an admirable life, then he has an argument that the traditional picture of the moral demand needs to be rejected. Only fools and bigots will be able to stay committed to morality in its traditional form.

Here I think it is relevant to look at the lives of those who have experienced great evil, and have yet persevered in their faith. What do these lives show us? That the experts in the experience of evil have not always found that this evidence forces them to reject their faith. I talked with a woman named Eva while I was writing *The Moral Gap*. She was a survivor of the German concentration camps of World War II. She said that her experience was that those who went into the camps with a strong faith in God came out, if they came out at all, with their faith stronger. It is not that they understood why God permitted the suffering, but that their faith in him is what held them and kept them through it. Eva was Jewish, and I do not know whether she believed in an afterlife or not. My sense is that she did not. But she did have a basic attitude of trust that God was somehow in charge, and that the good was more fundamental in the world than the evil in it, and would in the end win. There is a large biographical and autobiographical literature here, and I am not claiming that it speaks with a single voice. But Elie Wiesel, to take one other example, says that he has been angry with God, and has not answered the question of why God allowed the Holocaust. But he did not lose his faith, and in fact claims to have become closer to God through his protest.[51] I am not discussing here the argument from evil against the existence of God. Rather, I am looking at the character of lives which contain moral faith even in the presence of overwhelming evil. It is important that the lives of people like Eva are admirable. Their lives carry conviction. They have a reading of the enormous evil which they experience. This is not merely logically consistent with their faith (though that is important to argue as well). My point is, rather, that this reading lies behind lives which are obviously praiseworthy, whose goodness it would be perverse to deny.

I will end with one instructive case, that of Ivan and Alyosha in Dostoyevsky's

Brothers Karamazov. Ivan tells the story of the Grand Inquisitor, which is often used in Western philosophy textbooks as the paradigm case of the argument against faith in providence. Ivan ends up, after telling a number of horrifying stories of evil, saying to God that he cannot believe in a providential order, that he returns the ticket. But we have to ask what Dostoyevsky is doing in putting this powerful section in the mouth of Ivan, with Alyosha as its audience. The philosophy textbooks take the passage out of its context in the whole novel, and they miss its point. What happens to these two brothers? Ivan, who does not disguise either the moral demand or his own radical incapacity, ends the book by going crazy and holding conversations with the devil. Alyosha ends the book declaring his faith to a circle of adoring children. Dostoyevsky is trying to show us something. The story of the Grand Inquisitor *is* powerful, but it is not decisive. The life of Ivan, who is a man in the moral gap but without moral faith, is doomed; but Alyosha, who retains his faith without pretending that evil does not exist, is bound for glory. We can see in his life the character of the full kingdom of God towards which he is headed.

What is the implication of all this for the relation between naturalism and morality? We can put this both negatively and positively. Negatively, the point is that there is an incoherence involved in accepting both naturalism and the moral-gap picture of morality. Positively, the point is that we need moral faith if we are to sustain moral lives through the experience of evil. If naturalists have a way to give us such moral faith, they need to present it to us.

Notes

1 I am using the term "naturalism" here as equivalent to the denial of the supernatural. The editors of this volume use the term "philosophical naturalism" for this purpose. A naturalist in this sense is someone who denies there is anything outside or beyond (anything which transcends) this universe. Unfortunately, the term "naturalism" also has a particular use within moral philosophy which is not relevant here, but which needs to be explicitly set aside. One place to find a reasonably precise definition for this use is in R. M. Hare's *Sorting Out Ethics* (Oxford: Clarendon Press, 1997), 63–81. This definition is embedded in his own theory, and is thus somewhat technical. Readers who do not know his vocabulary can ignore the present footnote, since it will not affect the rest of the paper. R. M. Hare divides up ethical theories first into descriptivist and non-descriptivist theories, where the difference between these is that descriptivist theories affirm, but non-descriptivist theories deny, that apart from syntactical features, the meanings of moral statements are determined entirely by their truth conditions. He then divides descriptivist theories by the kind of truth conditions that they say determine the meanings of moral statements. *Naturalist* theories hold that the truth conditions are the possession (by the actions, people, etc. about which the moral statements are made) of what he calls, following tradition, "natural properties." Intuitionist theories hold that the properties in question are "non-natural" or "*sui generis*." The difference here is that a naturalist is someone who thinks that the truth conditions of moral statements require the possession by their subjects of properties which can be defined, or have their meaning

explained, without using any specifically moral terms. An intuitionist is someone who thinks that in order to give the truth conditions of these statements we have to use specifically moral terms. In this sense of "naturalist" it is possible to avoid naturalism by being an intuitionist, and also by avoiding descriptivism of both these kinds (as R. M. Hare himself does). I mention this sense of the term, however, only in order to set it aside. In the sense of "naturalism" I am using in this paper, R. M. Hare is a naturalist, as are most intuitionists. See, for example, R. M. Hare, "The Simple Believer," in *Essays on Religion and Education* (Oxford: Clarendon Press, 1992), 24–5.

2 The argument I am giving in this paper has been given more fully in my book, *The Moral Gap* (Oxford: Clarendon Press, 1996). The reader will find many of the points I make in the present paper more adequately covered in this book, where I also engage more explicitly with the work of other philosophers. I have had to omit in the present essay all sorts of refinements and qualifications, and the result is that the thesis appears in a stark form. This is in some ways an improvement. I have been able in some places to put the argument more clearly just because of the constraints of the project.

3 Immanuel Kant, *Critique of Practical Reason*, 5: 130, and *Religion within the Limits of Reason Alone*, 6: 154. I will refer to Kant by the volume and page numbers of the Prussian Academy Edition, except for his *Lectures on Ethics*, where I will refer to the English translation. I have tried to explain what Kant means by saying we should see our duties as God's command, in "Kant's divine command theory," forthcoming in *Kant and Kierkegaard on Religion*, ed. D. Z. Phillips (New York: Westview, 2000).

4 Kant, *Groundwork of the Metaphysics of Morals*, 4: 421.

5 Ibid., 4: 429.

6 I discuss this question in *The Moral Gap*, chapter 6.

7 John Stuart Mill, *Utilitarianism* (Indianapolis: Hackett, 1979), 4.

8 See, R.M. Hare, *The Language of Morals* (Oxford: Clarendon Press, 1952); John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971); Jurgen Habermas, *The Theory of Communicative Action* (Boston: Beacon Press, 1984, 1987).

9 Kant, *Religion within the Limits of Reason Alone*, 6:26 and 6:29.

10 Kant also believed in moral progress through education. See *Education*, 9: 446. But he believed (though this is not fully worked out until *Religion within the limits of Reason Alone*, see 6: 98) that this change over the history of humankind, like the change within an individual, required divine assistance.

11 Henry Bedinger Mitchell, *Talks on Religion* (New York: Longmans, Green, and Co., 1908), 47.

12 Mill, *Utilitarianism*, 17.

13 R. M. Hare, *Moral Thinking* (Oxford: Clarendon Press, 1981), 99. See also p. 34.

14 Friedrich Nietzsche, *On the Genealogy of Morals* and *Ecce Homo*, trans. Walter Kaufman (New York: Vintage Books, 1967), 90–1.

15 For example, Samuel Scheffler, *Human Morality* (Oxford: Oxford University Press, 1992), and Bernard Williams, *Shame and Necessity* (Berkeley: University of California Press, 1994).

16 Nietzsche is illuminating for another reason. He has his own three-part structure. His positive doctrine is hard to discern, but it seems to contain a demand which can be met by the overman, but not by the rest of us. But to what extent, then, can we be held accountable to this demand?

17 Aristotle, *Nicomachean Ethics*, 10. 7, 1177b 26–34.

18 Fung Yu-Lan, *A Short History of Chinese Philosophy* (New York: The Free Press, 1966; reprinted

in 1977), 302. Chu Hsi adds that Heaven is the source of principle, or Li.

19 Augustine, *On Grace and Free Will*, 16. 32.

20 Luther, *The Bondage of the Will*, trans. J. J. Packer and O. R. Johnston (Fleming H. Revell Co. 1957), 152.

21 This is why Kierkegaard's pastor at the end of *Either/Or* introduces the thought, "you might actually have done what you could but no one came to your assistance." Kierkegaard, *Either/Or*, volume 2, ed. and trans. Howard H. Hong and Edna H. Hong (Princeton: Princeton University Press, 1987), 345. It is this kind of thought about our capacities, indexed to the available assistance, that leads in Kierkegaard's system out of the ethical life and into the religious.

22 *The Moral Gap* deals with both the description and the evaluation of these strategies at greater length in chapters 4–7.

23 Shelly Kagan, *The Limits of Morality* (Oxford: Clarendon Press, 1989).

24 John E. Hare and Carey B. Joynt, *Ethics and International Affairs* (London: Macmillan, 1982), 163–83.

25 Judith Jarvis Thomson, "A Defense of Abortion" *Philosophy and Public Affairs*, vol. 1, no. 1, 47–66, and then endlessly reprinted in anthologies.

26 James Griffin, *Well-Being* (Oxford: Clarendon Press, 1986), 114.

27 Germain C. Grisez, "The First Principle of Practical Reason: A Commentary on *Summa Theologiae*," 1–2, Question 94, Article 2, *in Aquinas: A Collection of Critical Essays*, ed. Anthony Kenny (South Bend, Ind.: Notre Dame Press, 1969), 340–82. Joseph M. Boyle, "Aquinas and Prescriptive Ethics," *Proceedings of the American Catholic Philosophical Association* 49 (1975): 82–95. John Finnis, *Natural Law and Natural Rights*, (Oxford: Clarendon Press, 1980), chapter 4. But these authors do not try to *deduce* how we ought to act from what we in fact aim at.

28 It is interesting to observe that Aristotle *does* include power and prestige as requirements of the human good. See *Nicomachean Ethics* I, 2, 1094b8–10; IX, 7, 1168a5–10; VI, 5, 1140b8–11; VI, 8, 1141b24; *Politics* III, 4, 1277a13–b29. This is one main reason why contemporary theorists have felt nervous about embracing an Aristotelian view too closely.

29 Kant, *Religion within the Limits of Reason Alone*, 6: 93.

30 Nel Noddings, *Caring: A Feminine Approach to Ethics and Moral Education* (Berkeley: University of California Press, 1984), 86.

31 In *The Moral Gap*, chapter 6, I argue that there is a part of morality that does not conform to Kant's formula of universal law.

32 If Noddings wants to include this under "crossing our path," then her account is not different any longer from the universal caring she wants to deny.

33 The parent is also a person, and the child needs to learn this; and the parent needs to learn that certain forms of self-sacrifice for the sake of the child are illegitimate. Again, however, this argument begs the question. An extreme particularist could deny that special relations are morally limited in this way.

34 Donald T. Campbell, "On the Conflicts between Biological and Social Evolution and between Psychology and Moral Tradition, "*American Psychologist*, 30 (1975): 1103–26. This was his presidential address to the American Psychological Association.

35 Donald T. Campbell, "Legal and Primary Group Social Controls,"*Journal of Social and Biological Structures*, 5 (1982): 431–8.

36 Allan Gibbard, *Wise Choices, Apt Feelings* (Cambridge, Mass.: Harvard University Press, 1990). Gibbard is better than other evolutionary ethicists, in that he recognizes that charity or

altruism "may have no apparent reproductive payoff" (Ibid., 260).

37 Gibbard, *Wise Choices*, 140.

38 For example, Allen Wood, "Kant's Deism," in *Kant's Philosophy of Religion Reconsidered*, ed. Philip J. Rossi and Michael Wreen (Bloomington, Ind.: Indiana University Press, 1991), 1–21.

39 Kant, *Religion within the Limits of Reason Alone*, 6: 155. I argue for the attribution in *The Moral Gap*, chapter 2.

40 *Conflict of the Faculties*, 7: 54. For the limitation of theoretical and practical reason here, see *Religion within the Limits of Reason Alone*, 6: 53, and *Critique of Judgement*, 5: 176.

41 Kant, *Conflict of the Faculties*, 6: 54.

42 Kant, *Lectures on Ethics*, trans. Louis Infield (Indianapolis: Hackett, 1930), 91–3.

43 Kant, *Religion within the Limits of Reason Alone*, 6: 67.

44 Ibid., 6: 60.

45 Kant, *Critique of Practical Reason*, 5: 22.

46 This is the first kind of moral faith I distinguished earlier, faith in the possibility of one's virtue. The structure of this subsection of the paper is complex. I have distinguished two kinds of moral faith, and then distinguished the second of these (the faith in Providence) into a less and a more ambitious belief, and then I have claimed to derive the first of these from two more familiar moral beliefs, the first of which is the first kind of moral faith I started with.

47 Actually the relevant belief is more specific than this. I think the moral agent has to believe that there is the right kind of causal chain between her willing the good and her achieving it a significant proportion of the time. But there are many complexities here about what kind of causal chain is involved, and I will not go into them.

48 We can add another point in the context of the present argument. If everyone were *virtuous*, the main obstacle to peoples' accomplishing their good aims would be removed, namely the vicious interference of others.

49 Kant, *Critique of Pure Reason*, A810 = B838

50 Bernard Williams, *Shame and Necessity*, (Berkeley: University of California Press, 1994), 163–4.

51 *New York Times*, 15 October 1986, sec. A, 1, 10.

Part 4

Natural theology

9 Naturalism and cosmology

William Lane Craig

The fundamental question

In his biography of Ludwig Wittgenstein, Norman Malcolm reports,

> He said that he sometimes had a certain experience which could best be described by saying that 'when I have it, *I wonder at the existence of the world.* I am then inclined to use such phrases as "How extraordinary that anything should exist!" or "How extraordinary that the world should exist!"'[1]

This mystery, which according to Aristotle lay at the very root of philosophy,[2] is one which thoughtful naturalists cannot avoid. Derek Parfit, for example, agrees that "No question is more sublime than why there is a Universe: why there is anything rather than nothing."[3]

As we all know, this question led G. W. Leibniz to posit the existence of a necessary being, which carries within itself the sufficient reason for its existence and which constitutes the sufficient reason for the existence of everything else in the world.[4] Leibniz identified this being as God. Naturalists, on the other hand, have typically claimed that the space-time universe is itself at least factually necessary – that is to say, eternal, uncaused, incorruptible, and indestructible[5] – while dismissing the demand for a logically necessary being. Thus, David Hume queried, "Why may not the material universe be the necessarily existent Being . . . ?"[6] Indeed, "How can anything, that exists from eternity, have a cause, since that relation implies a priority in time and a beginning of existence?"[7] There is no warrant for going beyond the universe to posit a supernatural ground of its existence. As Bertrand Russell put it so succinctly in his BBC radio debate with Frederick Copleston, "The universe is just there, and that's all."[8]

The origin of the universe

It was thus the presumed eternity of the universe that allowed naturalistic minds to rest comfortably in the face of the mystery of existence. This feeling of *Gemütlichkeit* was

first disturbed when, in 1917, Albert Einstein made a cosmological application of his newly discovered gravitational theory, the General Theory of Relativity.[9] In so doing he assumed that the universe is homogeneous and isotropic and that it exists in a steady state, with a constant mean mass density and a constant curvature of space. To his chagrin, however, he found that GR would not permit such a model of the universe unless he introduced into his gravitational field equations a certain "fudge factor" in order to counterbalance the gravitational effect of matter and so ensure a static universe. But Einstein's universe was balanced on a razor's edge, and the least perturbation would cause the universe either to implode or to expand. By taking this feature of Einstein's model seriously, the Russian mathematician Alexander Friedman and the Belgian astronomer Georges Lemaître were able to formulate independently in the 1920s solutions to the field equations which predicted an expanding universe.[10]

The monumental significance of the Friedman–Lemaître model lay in its historization of the universe. As one commentator has remarked, up to this time the idea of the expansion of the universe "was absolutely beyond comprehension. Throughout all of human history the universe was regarded as fixed and immutable and the idea that it might actually be changing was inconceivable."[11] But if the Friedman–Lemaître model were correct, the universe could no longer be adequately treated as a static entity existing, in effect, timelessly. Rather the universe has a history, and time will not be a matter of indifference for our investigation of the cosmos.

In 1929 the astronomer Edwin Hubble showed that the red-shift in the optical spectra of light from distant galaxies was a common feature of all measured galaxies and was proportional to their distance from us.[12] This red-shift, first observed by Vesto Slipher is 1926, was taken to be a Doppler effect indicative of the recessional motion of the light source in the line of sight. Incredibly, what Hubble had discovered was the isotropic expansion of the universe predicted by Friedman and Lemaître on the basis of Einstein's GR. It was a veritable turning point in the history of science. "Of all the great predictions that science has ever made over the centuries," exclaims John Wheeler, "was there ever one greater than this, to predict, and predict correctly, and predict against all expectation a phenomenon so fantastic as the expansion of the universe?"[13]

The Standard Big Bang model

According to the Friedman–Lemaître model, as time proceeds, the distances separating material particles become greater. It is important to understand that as a GR-based theory, the model does not describe the expansion of the material content of the universe into a pre-existing, empty, Newtonian space, but rather the expansion of space itself. The ideal particles of the cosmological fluid constituted by the matter and energy of the universe are conceived to be at rest with respect to space but to recede progressively from one another as space itself expands or stretches, just as buttons glued to the surface of a balloon would recede from one another as the

balloon inflates. As the universe expands, it becomes less and less dense. This has the astonishing implication that as one reverses the expansion and extrapolates back in time, the universe becomes progressively denser until one arrives at a state of infinite density at some point in the finite past. This state represents a singularity at which space-time curvature, along with temperature, pressure and density, becomes infinite. It therefore constitutes an edge or boundary to space-time itself. P. C. W. Davies comments,

> If we extrapolate this prediction to its extreme, we reach a point when all distances in the universe have shrunk to zero. . . An initial cosmological singularity therefore forms a past temporal extremity to the universe. We cannot continue physical reasoning, or even the concept of spacetime, through such an extremity. For this reason most cosmologists think of the initial singularity as the "beginning" of the universe. On this view the big bang represents the creation event; the creation not only of all the matter and energy in the universe, but also of spacetime itself.[14]

The term "Big Bang," originally a derisive expression coined by Fred Hoyle to characterize the beginning of the universe predicted by the Friedman– Lemaître model, is thus potentially misleading, since the expansion cannot be visualized from the outside (there being no "outside," just as there is no "before" with respect to the Big Bang).[15]

The Standard Big Bang model, as the Friedman–Lemaître model came to be called, thus describes a universe which is not eternal in the past, but which came into being a finite time ago. Moreover – and this deserves underscoring – the origin it posits is an absolute origin *ex nihilo*. For not only all matter and energy, but space and time themselves come into being at the initial cosmological singularity. As Barrow and Tipler emphasize, "At this singularity, space and time came into existence; literally nothing existed before the singularity, so, if the Universe originated at such a singularity, we would truly have a creation *ex nihilo*."[16] Thus, we may graphically represent space-time as a cone (Figure 9.1).

On such a model the universe originates *ex nihilo* in the sense that at the initial singularity it is true that *There is no earlier space-time point* or it is false that *Something existed prior to the singularity*.

Now such a conclusion is profoundly disturbing for anyone who ponders it. For the question cannot be suppressed: *Why does the universe exist rather than nothing?* Sir Arthur Eddington, contemplating the beginning of the universe, opined that the expansion of the universe was so preposterous and incredible that "I feel almost an indignation that anyone should believe in it – except myself."[17] He finally felt forced to conclude, "The beginning seems to present insuperable difficulties unless we agree to look on it as frankly supernatural."[18]

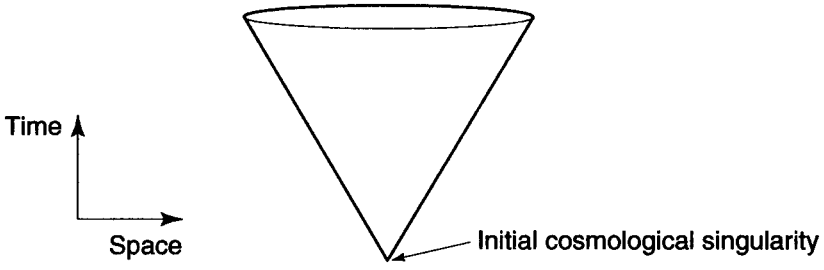

Figure 9.1 Conical representation of Standard model space-time. Space and time begin at the initial cosmological singularity, before which literally nothing exists.

Standard Big Bang cosmogony thus presents a challenge to scientific naturalism, since, in Quentin Smith's words, "It belongs analytically to the concept of the cosmological singularity that it is not the effect of prior physical events. The definition of a singularity . . . entails that it is *impossible to extend the spacetime manifold beyond the singularity* . . . . This rules out the idea that the singularity is an effect of some prior natural process."[19] Smith recognizes that the question which then remains is whether the Big Bang might not be plausibly regarded as the result of a supernatural cause. Otherwise, one must say that the universe simply sprang into being uncaused out of absolutely nothing. Thus, in the words of one astrophysical team, "The problem of the origin involves a certain metaphysical aspect which may be either appealing or revolting."[20]

The Steady State model

Revolted by the stark metaphysical alternatives presented by an absolute beginning of the universe, naturalists have been understandably eager to subvert the Standard model and restore an eternal universe. Sir Fred Hoyle, for example, could countenance neither an uncaused nor a supernaturally caused origin of the universe. With respect to the singularity, he wrote, "This most peculiar situation is taken by many astronomers to represent *the origin of the universe*. The universe is supposed to have begun at this particular time. From where? The usual answer, surely an unsatisfactory one, is: from nothing!"[21] Equally unsatisfactory was the postulation of a supernatural cause. Noting that some accept happily the universe's absolute beginning, Hoyle complained,

> To many people this thought process seems highly satisfactory because a "something" outside physics can then be introduced at $\tau = 0$. By a semantic manoeuvre, the word "something" is then replaced by "god," except that the first letter becomes a capital, God, in order to warn us that we must not carry the enquiry any further.[22]

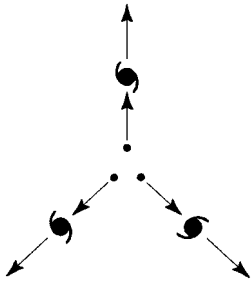

Figure 9.2 Steady State model. As the galaxies mutually recede, new matter comes into existence to replace them. The universe thus constantly renews itself and so never began to exist.

To Hoyle's credit, he did carry the inquiry further by helping to formulate the first competitor to the Standard model. In 1948 Hoyle, together with Hermann Bondi and Thomas Gold, broached the Steady State model of the universe.[23] According to this theory, the universe is in a state of isotropic cosmic expansion, but as the galaxies recede, new matter is drawn into being *ex nihilo* in the interstices of space created by the galactic recession (Figure 9.2).

The expansion of the universe in the Steady State model can be compared to a rubber sheet with buttons glued to it: as the sheet is stretched and the buttons separate, new buttons come into being in the voids created by the recession of the previously existing buttons. Thus, the condition of the sheet remains constant over time, and no beginning of the process need be posited. If one extrapolates the expansion of the universe back in time, the density of the universe never increases because the matter and energy simply vanish as the galaxies mutually approach!

The Steady State theory never secured a single piece of experimental verification; its appeal was purely metaphysical.[24] The discovery of progressively more radio galaxies at ever greater distances undermined the theory by showing that in the past the universe was significantly different than it is today, thus contradicting the notion of the steady state of the universe. Instead it became increasingly evident that the universe had an evolutionary history. But the decisive refutation of the Steady State model came with two discoveries which constituted, in addition to the galactic red-shift, the most significant evidence for the Big Bang theory: the cosmogonic nucleosynthesis of the light elements and the microwave background radiation. Although the heavy elements were synthesized in the stellar furnaces, stellar nucleosynthesis could not manufacture the abundant light elements such as helium and deuterium. These could only have been created in the extreme conditions present in the first moment of the Big Bang. In 1965 a serendipitous discovery revealed the existence of a cosmic background radiation predicted in the 1940s by

George Gamow on the basis of the Standard model. This radiation, now shifted into the microwave region of the spectrum, consists of photons emitted during a very hot and dense phase of the universe. In the minds of almost all cosmologists, the cosmic background radiation decisively discredited the Steady State model.

Oscillating models

The Standard model was based on the assumptions of homogeneity and isotropy. In the 1960s and 1970s some cosmologists suggested that by denying homogeneity and isotropy, one might be able to craft an Oscillating model of the universe and thus avert the absolute beginning predicted by the Standard model.[25] If the internal gravitational pull of the mass of the universe were able to overcome the force of its expansion, then the expansion could be reversed into a cosmic contraction, a Big Crunch. If the universe were not homogeneous and isotropic, then the collapsing universe might not coalesce at a point, but the material contents of the universe might pass by one another, so that the universe would appear to bounce back from the contraction into a new expansion phase. If this process could be repeated indefinitely, then an absolute beginning of the universe might be avoided (Figure 9.3).

Such a theory is extraordinarily speculative, but again there were metaphysical motivations for adopting this model.[26] The prospects of the Oscillating model were severely dimmed in 1970, however, by Roger Penrose and Stephen Hawking's formulation of the Singularity Theorems which bear their names.[27] The theorems disclosed that under very generalized conditions an initial cosmological singularity is inevitable, even for inhomogeneous and non-isotropic universes. Reflecting on the impact of this discovery, Hawking notes that the Hawking–Penrose Singularity Theorems "led to the abandonment of attempts (mainly by the Russians) to argue that there was a previous contracting phase and a non-singular bounce into expansion. Instead almost everyone now believes that the universe, and time itself, had a beginning at the big bang."[28]

Despite the fact that no space-time trajectory can be extended through a singularity, the Oscillating model exhibited a stubborn persistence. Three further strikes were lodged against it. First, there are no known physics which would cause a collapsing universe to bounce back to a new expansion. If, in defiance of the Hawking–Penrose Singularity Theorems, the universe rebounds, this is predicated upon a physics which is completely unknown. Physics predicts that a universe in a state of gravitational self-collapse will not rebound like a basketball dropped to the floor, but rather land like a lump of clay.[29] Second, the observational evidence indicates that the mean mass density of the universe is insufficient to generate enough gravitational attraction to halt and reverse the expansion. Tests employing a variety of techniques for measuring the density of the universe and the deceleration of the expansion continue to point to a density of 0.1–0.2 the critical value. Most recently, in January of 1998, astronomical teams from Princeton, Yale, the Lawrence Berkeley National Laboratory, and the Harvard-Smithsonian Astrophysics Institute reported at the

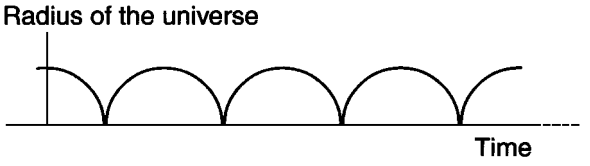

Figure 9.3 Oscillating model. Each expansion phase is preceded and succeeded by a contraction phase, so that the universe in concertina-like fashion exists beginninglessly and endlessly.

American Astronomical Society meeting that their various tests all show that "the universe will expand forever."[30] A spokesman for the Harvard–Smithsonian team stated that they were now at least 95 per cent certain that "the density of matter is insufficient to halt the expansion of the universe."[31] This effectively rules out an oscillating universe. Third, the thermodynamic properties of an Oscillating model imply the very beginning its proponents sought to avoid. Entropy increases from cycle to cycle in such a model, which has the effect of generating larger and longer oscillations with each successive cycle (Figure 9.4). Thus, as one traces the oscillations back in time, they become progressively smaller until one reaches a first and smallest oscillation. Zeldovich and Novikov conclude, "The multicycle model has an infinite future, but only a finite past."[32] In fact, astronomer Joseph Silk estimates on the basis of current entropy levels that the universe cannot have gone through more than 100 previous oscillations.[33] Even if this difficulty were avoided,[34] a universe oscillating from eternity past would require an infinitely precise tuning of initial conditions in order to perdure through an infinite number of successive bounces. A universe rebounding from a single, infinitely long contraction is, if entropy increases during the contracting phase, thermodynamically untenable and incompatible with the initial low entropy condition of our expanding phase; postulating an entropy decrease during the contracting phase in order to escape this problem would require us to postulate inexplicably special low entropy conditions at the time of the bounce in the life of an infinitely evolving universe. Such a low entropy condition at the beginning of the expansion is more plausibly accounted for by the presence of a singularity or some sort of quantum creation event.

Although these difficulties were well known, proponents of the Oscillating model tenaciously clung to it until a new alternative to the Standard model emerged during the 1970s.[35] Looking back, quantum cosmologist Christopher Isham muses,

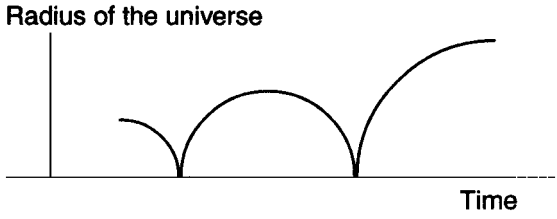

Figure 9.4 Oscillating model with entropy increase. Due to the conservation of entropy each successive oscillation has a larger radius and longer expansion time.

> Perhaps the best argument in favor of the thesis that the Big Bang supports theism is the obvious unease with which it is greeted by some atheist physicists. At times this has led to scientific ideas, such as continuous creation or an oscillating universe, being advanced with a tenacity which so exceeds their intrinsic worth that one can only suspect the operation of psychological forces lying very much deeper than the usual academic desire of a theorist to support his/her theory.[36]

The Oscillating model drew its life from its avoidance of an absolute beginning of the universe; but once other models became available claiming to offer the same benefit, the Oscillating model sank into oblivion under the weight of its own deficiencies.

Vacuum Fluctuation models

It was realized that a physical description of the universe prior to the Planck time (10^{-43} second after the Big Bang singularity) would require the introduction of quantum physics in addition to GR. On the quantum level, so-called virtual particles are thought to arise due to fluctuations in the energy locked up in the vacuum, particles which the Heisenberg Indeterminacy Principle allows to exist for a fleeting moment before dissolving back into the vacuum. In 1973 Edward Tryon speculated whether the universe might not be a long-lived virtual particle, whose total energy is zero, born out of the primordial vacuum.[37] This seemingly bizarre speculation gave rise to a new generation of cosmogonic theories which we may call Vacuum Fluctuation models. These models were closely related to an adjustment to the Standard model known as Inflation. In an attempt to explain – or explain away, depending on one's viewpoint – the astonishing large-scale homogeneity and isotropy of the universe, certain theorists proposed that between 10^{-35} and 10^{-33} sec after the Big Bang singularity, the universe underwent a phase of super-rapid, or inflationary, expansion which served to push the

inhomogeneities out beyond our event horizon.[38] Prior to the Inflationary Era the universe was merely empty space, or a vacuum, and the material universe was born when the vacuum energy was converted into matter via a quantum mechanical phase transition. In most inflationary models, as one extrapolates backward in time, beyond the Planck time, the universe continues to shrink down to the initial singularity. But in Vacuum Fluctuation models, it is hypothesized that prior to inflation the Universe-as-a-whole was not expanding. This Universe-as-a-whole is a primordial vacuum which exists eternally in a steady state. Throughout this vacuum sub-atomic energy fluctuations constantly occur, by means of which matter is created and mini-universes are born (Figure 9.5). Our expanding universe is but one of an indefinite number of mini-universes conceived within the womb of the greater Universe-as-a-whole. Thus, the beginning of our universe does not represent an absolute beginning, but merely a change in the eternal, uncaused Universe-as-a-whole.

Though still bandied about in the popular press, Vacuum Fluctuation models did not outlive the decade of the 1980s. Not only were there theoretical problems with the production mechanisms of matter, but these models faced a deep internal incoherence.[39] According to such models, it is impossible to specify precisely when and where a fluctuation will occur in the primordial vacuum which will then grow into a universe. Within any finite interval of time there is a positive probability of such a fluctuation occurring at any point in space. Thus, given infinite past time, universes will eventually be spawned at *every* point in the primordial vacuum, and, as they expand, they will begin to collide and coalesce with one another. Thus, given infinite past time, we should by now be observing an infinitely old universe, not a relatively young one. One theorist tries to avoid this problem by stipulating that fluctuations in the primordial vacuum only occur infinitely far apart, so that each mini-universe has infinite room in which to expand.[40] Not only is such a scenario unacceptably *ad hoc*, but it does not even solve the problem. For given infinite past time, each of the infinite regions of the vacuum will have spawned an open universe which by now will have entirely filled that region, with the result that all of the individual mini-universes would have coalesced.

Isham has called this problem "fairly lethal" to Vacuum Fluctuation models and says that therefore they "have not found wide acceptance."[41] About the only way to avert the problem would be to postulate an expansion of the primordial vacuum itself; but then we are right back to the absolute origin implied by the Standard model. According to Isham these models were therefore "jettisoned twenty years ago" and "nothing much" has been done with them since.[42]

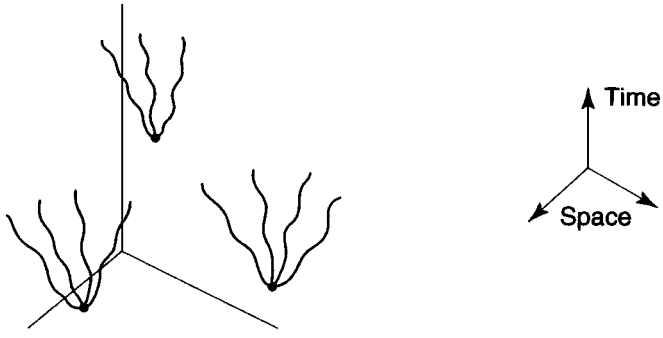

Figure 9.5 Vacuum Fluctuation models. Within the vacuum of the wider universe, fluctuations occur which grow into mini-universes. Ours is but one of these, and its relative beginning does not imply a beginning for the Universe-as-a-whole.

Chaotic Inflationary model

Inflation also forms the context for the next alternative we shall consider: the Chaotic Inflationary model. Inflationary theory has not only been criticized as unduly "metaphysical," but has also been crippled by various physical problems (such as getting inflation to transition to the current expansion). We have seen the Old Inflationary Theory and the New Inflationary Theory, both of which are now dead. One of the most fertile of the inflation theorists has been the Russian cosmologist Andrei Linde, who currently champions his Chaotic Inflationary model.[43] According to cosmologist Robert Brandenberger, "Linde's chaotic inflation scenario is . . . the only viable inflationary model in the sense that it is not plagued with internal inconsistencies (as 'old inflation' and 'new inflation' are)."[44] In Linde's model, inflation *never* ends: each inflating domain of the universe when it reaches a certain volume gives rise via inflation to another domain, and so on, *ad infinitum* (Figure 9.6).

Linde's model thus has an infinite future. But Linde is troubled at the prospect of an absolute beginning. He writes, "The most difficult aspect of this problem is not the existence of the singularity itself, but the question of what was *before* the singularity . . . . This problem lies somewhere at the boundary between physics and metaphysics."[45] Linde therefore proposes that chaotic inflation is not only endless, but beginningless. Every domain in the universe is the product of inflation in another domain, so that the singularity is averted and with it as well the question of what came before (or, more accurately, what caused it).

In 1994, however, Arvind Borde and Alexander Vilenkin showed that a universe eternally inflating toward the future cannot be geodesically complete in the past; that is to say, there must have existed at some point in the indefinite past an initial singularity. They write,

Figure 9.6 Chaotic Inflationary model. The wider universe produces via inflation separate domains which continue to recede from one another as the wider space expands.

> A model in which the inflationary phase has no end . . . naturally leads to this question: Can this model also be extended to the infinite past, avoiding in this way the problem of the initial singularity?
>
> . . . this is in fact not possible in future-eternal inflationary space-times as long as they obey some reasonable physical conditions: such models must necessarily possess initial singularities.
>
> . . . the fact that inflationary spacetimes are past incomplete forces one to address the question of what, if anything, came before.[46]

In his response, Linde concurs with the conclusion of Borde and Vilenkin: there must have been a Big Bang singularity at some point in the past.[47]

Quantum Gravity models

At the close of their analysis of Linde's Chaotic Inflationary model, Borde and Vilenkin say with respect to Linde's metaphysical question, "The most promising way to deal with this problem is probably to treat the Universe quantum mechanically and describe it by a wave function rather than by a classical space time."[48] They thereby allude to the last class of models which we shall discuss, namely, Quantum Gravity models. Vilenkin and, more famously, James Hartle and Stephen Hawking have proposed models of the universe which Vilenkin candidly calls exercises in "metaphysical cosmology."[49] In his best-selling popularization of his theory, Hawking even reveals an explicitly theological orientation. He concedes that on the Standard model one could legitimately identify the Big Bang singularity as the instant at which God created the universe.[50] Indeed, he thinks that a number of attempts to avoid the Big Bang were probably motivated by the feeling that a beginning of time "smacks of divine intervention."[51] He sees his own model as preferable to the Standard model because there would be no edge of space-time at which one "would have to appeal to

God or some new law."[52] As we shall see, he is not at all reluctant to draw theological conclusions on the basis of his model.

Both the Hartle–Hawking and the Vilenkin models eliminate the initial singularity by transforming the conical hyper-surface of classical space-time into a smooth, curved hyper-surface having no edge (Figure 9.7).

This is accomplished by the introduction of imaginary numbers for the time variable in Einstein's gravitational equations, which effectively eliminates the singularity. Hawking sees profound theological implications in the model:

> The idea that space and time may form a closed surface without boundary . . . has profound implications for the role of God in the affairs of the universe . . . . So long as the universe had a beginning, we could suppose it had a creator. But if the universe is really completely self-contained, having no boundary or edge, it would have neither beginning nor end. What place, then, for a creator?[53]

Hawking does not deny the existence of God, but he does think his model eliminates the need for a Creator of the universe.

The key to assessing this theological claim is the physical interpretation of Quantum Gravity models. By positing a finite (imaginary) time on a closed surface prior the Planck time rather than an infinite time on an open surface, such models actually seem to support, rather than undercut, the idea that time had a beginning. Such theories, if successful, enable us to model the beginning of the universe without an initial singularity involving infinite density, temperature, pressure, and so on. As Barrow points out, "This type of quantum universe has not always existed; it comes into being just as the classical cosmologies could, but it does not start at a Big Bang where physical quantities are infinite . . . ,"[54] Barrow points out that such models are "often described as giving a picture of 'creation out of nothing,'" the only caveat being that in this case "there is no definite . . . point of creation."[55] Hartle–Hawking themselves construe their model as giving "the amplitude for the Universe to appear from nothing," and Hawking has asserted that according to the model the universe "would quite literally be created out of nothing: not just out of the vacuum, but out of absolutely nothing at all, because there is nothing outside the universe."[56] Similarly, Vilenkin claims that his model describes the creation of the universe "from literally *nothing*."[57] Taken at face value, these statements entail the beginning of the universe. Hence, Hawking means to include himself when he asserts, "Today, virtually everybody believes that the universe, and time itself, had a beginning at the Big Bang."[58] Hawking's statement quoted above concerning the theological implications of his model must therefore be understood to mean that on such models there are no beginning or ending *points*. But having a beginning does not entail having a beginning point. Even in the Standard model, theorists sometimes "cut out" the initial singular point without thinking that therefore space-time no longer begins to exist and the problem of the origin of the universe is thereby resolved. Time begins to exist just in

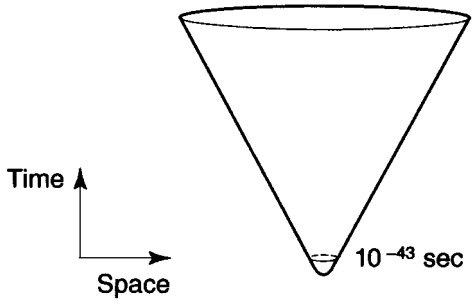

Figure 9.7 Quantum Gravity model. In the Hartle–Hawking version, space-time is "rounded off" prior to the Planck time, so that although the past is finite, there is no edge or beginning point.

case for any finite temporal interval, there are only a finite number of equal temporal intervals earlier than it. That condition is fulfilled for Quantum Gravity models as well as for the Standard model. Nor should we think that by giving the amplitude for the universe to appear from nothing quantum cosmologists have eliminated the need for a Creator, for that probability is conditional upon several choices which only the Creator could make (such as selecting the wave function of the universe) and is dubiously applied to absolute nothingness.[59] Thus, Quantum Gravity models, like the Standard model, imply the beginning of the universe.

Perhaps it will be said that such an interpretation of Quantum Gravity models fails to take seriously the notion of "imaginary time." Introducing imaginary numbers for the time variable in Einstein's equation has the peculiar effect of making the time dimension indistinguishable from space. But in that case, the imaginary time regime prior to the Planck time is not a space-time at all, but a Euclidean four-dimensional space. Construed realistically, such a four-space would be evacuated of all temporal becoming and would simply exist timelessly. Thus, Vilenkin characterizes this regime as a "state in which all our basic notions of space, time, energy, entropy, etc. lose their meaning."[60] Hawking describes it as "completely self-contained and not affected by anything outside itself. It would be neither created nor destroyed. It would just BE."[61]

The question which arises for this construal of the model is whether such an interpretation is meant to be taken realistically or instrumentally. On this score, there can be little doubt that the use of imaginary quantities for time is a mere mathematical device without ontological significance. For, first, there is no intelligible physical interpretation of imaginary time on offer. What, for example, would it mean to speak of the lapse of an imaginary second or of a physical object's enduring through two imaginary minutes? Second, time is metaphysically distinct from space, its moments being ordered by an *earlier than* relation which does not similarly order points in space.

But this essential difference is obscured by imaginary time. Thus, "imaginary time" is most plausibly construed as a mathematical *Hilfsmittel*. Barrow observes, "physicists have often carried out this 'change time into space' procedure as a useful trick for doing certain problems in ordinary quantum mechanics, although they did not imagine that time was *really* like space. At the end of the calculation, they just swop [*sic*] back into the usual interpretation of there being one dimension of time and three . . . dimensions of . . . space."[62] In his model, Hawking simply declines to re-convert to real numbers. If we do, then the singularity re-appears. Hawking admits, "Only if we could picture the universe in terms of imaginary time would there be no singularities. . . . When one goes back to the real time in which we live, however, there will still appear to be singularities."[63] Hawking's model is thus a way of re-describing a universe with a singular beginning point in such a way that that singularity is transformed away; but such a re-description is not realist in character.

Remarkably, Hawking has recently stated explicitly that he interprets the Hartle–Hawking model non-realistically. He confesses, "I'm a positivist . . . I don't demand that a theory correspond to reality because I don't know what it is."[64] Still more extreme, "I take the positivist view point that a physical theory is just a mathematical model and that it is meaningless to ask whether it corresponds to reality."[65] In assessing the worth of a theory, "All I'm concerned with is that the theory should predict the results of measurements."[66] The clearest example of Hawking's instrumentalism is his combination of an electron quantum tunneling in Euclidean space (with time being imaginary) and an electron/positron pair accelerating away from each other in Minkowski space-time.[67] This analysis is directly analogous to the Hartle-Hawking cosmological model; and yet no one would construe particle pair creation as literally the result of an electron transitioning out of a timelessly existing four-space into our classical space-time. It is just an alternative description employing imaginary numbers rather than real numbers.

Significantly, the use of imaginary quantities for time is an inherent feature of *all* Quantum Gravity models.[68] This precludes their being construed realistically as accounts of the origin of the space-time universe in a timelessly existing four-space. Rather they are ways of modeling the real beginning of the universe *ex nihilo* in such a way as to not involve a singularity. What brought the universe into being remains unexplained on such accounts.

Moreover, we are not without positive reasons for affirming the reality of the singular origin of space-time postulated by the Standard model. John Barrow has rightly cautioned that "one should be wary of the fact that many of the studies of quantum cosmology are motivated by the desire to avoid an initial singularity of infinite density, so they tend to focus on quantum cosmologies that avoid a singularity at the expense of those that might contain one.[69] Noting the same tendency, Roger Penrose states, "I have gradually come around to the view that it is actually misguided to ask that the space-time singularities of classical relativity should disappear when

standard techniques of quantum (field) theory are applied to them."[70] For if the initial cosmological singularity is removed, then "we should have lost what seems to me to be the best chance we have of explaining the mystery of the second law of thermodynamics."[71] What Penrose has in mind is the remarkable fact that as one goes back in time the entropy of the universe steadily decreases. Just how unusual this is can be demonstrated by means of the Bekenstein–Hawking formula for the entropy of a stationary black hole. The total observed entropy of the universe is 10^{88}. Since there are around 10^{80} baryons in the universe, the observed entropy per baryon must be regarded as extremely small. By contrast in a collapsing universe the entropy would be 10^{123} near the end. Comparison of these two numbers reveals how absurdly small 10^{88} is compared to what it might have been. Thus, the structure of the Big Bang must have been severely constrained in order that thermodynamics as we know it should have arisen. So how is this special initial condition to be explained? According to Penrose, we need the initial cosmological singularity, conjoined with the Weyl Curvature Hypothesis, according to which initial singularities (as opposed to final singularities) must have vanishing Weyl curvature.[72] In standard models, the Big Bang does possess vanishing Weyl curvature. The geometrical constraints on the initial geometry have the effect of producing a state of very low entropy. So the entropy in the gravitational field starts at zero at the Big Bang and gradually increases through gravitational clumping. The Weyl Curvature Hypothesis thus has the time asymmetric character necessary to explain the second law. By contrast, the Hartle–Hawking model "is very far from being an explanation of the fact that past singularities have small Weyl curvature whereas future singularities have large Weyl curvature."[73] On Hawking's time symmetrical theory, we should have white holes spewing out material, in contradiction to the Weyl Curvature Hypothesis, the Second Law of Thermodynamics, and probably also observation.[74] Penrose supplies Figure 9.8 to illustrate the difference.

If we remove the initial cosmological singularity, we render the Weyl Curvature Hypothesis irrelevant and "we should be back where we were in our attempts to understand the origin of the second law."[75] Could the special initial geometry have arisen sheerly by chance in the absence of a cosmic singularity? Penrose's answer is decisive: "Had there not been any constraining principles (such as the Weyl curvature hypothesis) the Bekenstein–Hawking formula would tell as that the probability of such a 'special' geometry arising by chance is at least as small as about one part in $10^{1000B(3/2)}$ where B is the present baryon number of the universe [$\sim 10^{80}$]."[76] Thus Penrose calculates that, aiming at a phase space whose regions represent the likelihood of various possible configurations of the universe, "the accuracy of the Creator's aim" would have to have been one part in $10^{10(123)}$ in order for our universe to exist.[77] He comments, "I cannot even recall seeing anything else in physics whose accuracy is known to approach, even remotely, a figure like one part in $10^{10\,(123)}$."[78]

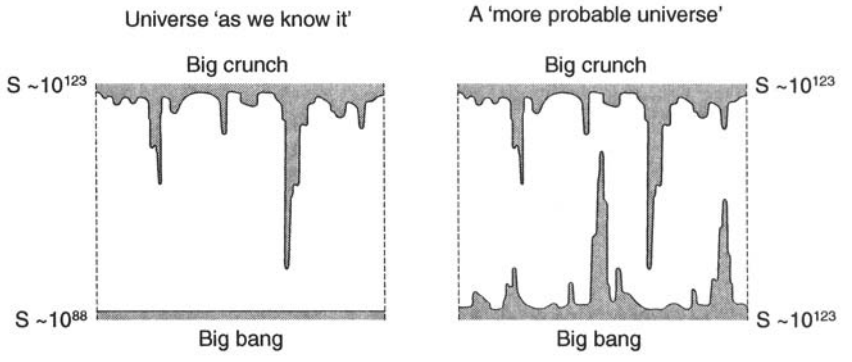

Figure 9.8 Contrast between the universe as we know it (assumed for convenience to be closed) with a more probable universe. In both cases the Big Crunch is a high entropy ($\sim 10^{123}$), complicated, unconstrained singularity. For the left-hand picture the Big Bang is a low entropy ($\leq 10^{88}$), highly constrained, initial singularity, while for the right-hand picture it is an unconstrained, much more probable Big Bang. The "stalactites" represent singularities of black holes, while the "stalagmites" represent singularities of white holes.

Thus, the initial cosmological singularity may be a virtual thermodynamical necessity. But whether it was at a singular point or not, the fact that the universe began to exist remains a prediction of any realistic interpretation of Quantum Gravity models.

Summary

With each successive failure of alternative cosmogonic theories to avoid the absolute beginning of the universe predicted by the Standard model, that prediction has been corroborated. This beginning of the universe, of space and time themselves, reveals the contingency of the universe. The universe is evidently not necessarily existent, as Hume suggested, since it is not eternal, and therefore its existence does cry out for explanation. It is no longer sufficient to dismiss this problem with a shrug and a slogan, "The universe is just there, and that's all."

Of course, in view of the metaphysical issues raised by the prospect of a beginning of the universe, we may be confident that the quest to avert such a beginning will continue unabated.[79] Such efforts are to be encouraged, and we have no reason to think that such attempts at falsification will result in anything other than further corroboration of the prediction of a beginning. In the meantime, the beginning cannot be wished away. Given its origin *ex nihilo*, the demand why the universe exists rather than nothing presses insistently upon us.

Beyond the Big Bang

The alternatives before us

The discovery that the universe is not eternal but had a beginning implies that the universe is not necessary in its existence and therefore has its ground in a transcendent being. The only way of avoiding this conclusion would be to deny Leibniz's conviction that whatever exists has a reason for its existence, either in the necessity of its own nature or else in an external ground, and to claim that the universe simply sprang into being uncaused out of nothing. Reflecting upon the current situation, P. C. W. Davies muses,

> What caused the big bang?' . . . One might consider some supernatural force, some agency beyond space and time as being responsible for the big bang, or one might prefer to regard the big bang as an event without a cause. It seems to me that we don't have too much choice. Either . . . something outside of the physical world . . . or . . . an event without a cause.[80]

In a remarkable article in the *Physical Review* of last year, J. Richard Gott and Li-Xin Li seek to break this dilemma by defending the extraordinary hypothesis that *the universe created itself*. Observing that "The question of first-cause has been troubling to philosophers and scientists alike for over two thousand years," they note that modern scientists have, like Aristotle, found models of the universe attractive which involve the universe's sempiternal existence, since in this way "one would not have to ask what caused it to come into being."[81] "Now that it is apparent that our universe began in a big bang explosion," however, "models with a finite beginning have taken precedence" and "the question of what happened before the big bang arises."[82] They observe that inflation seemed to be "a very promising answer, but as Borde and Vilenkin have shown, the inflationary state preceding the big bang could not have been infinite in duration – it must have had a beginning also. Where did it come from? Ultimately, the difficult question seems to be how to make something out of nothing."[83] Gott and Li-Xin, however, suggest instead that we should ask whether anything in the laws of physics would prevent the universe from creating itself.

Noting that General Relativity allows for the possibility of closed timelike curves, they hypothesize that as we trace the history of the universe back through an original inflationary state, we encounter a region of closed time-like curves prior to inflation. According to one possible scenario, a metastable vacuum inflates, producing an infinite number of (Big Bang type) bubble universes. In many of these a number of bubbles of metastable vacuum are created at late times by high energy events. These bubbles usually collapse and form black holes, but occasionally one will tunnel to

create an expanding, metastable vacuum or baby universe. One of these expanding, metastable vacuum baby universes "turns out to be the original inflating metastable vacuum we began with" (Figure 9.9). Gott and Li-Xin conclude that "the laws of physics may allow the universe to be its own mother."[84]

Now we may leave it to the physicists to assess Gott and Li-Xin's claim that the laws of physics permit such a scenario, as well as the question of whether there are non-law-like physical facts which contradict it. For the Gott-Li-Xin hypothesis raises fundamental metaphysical issues about the nature of time which, I think, render their hypothesis either metaphysically impossible or else superfluous.

Philosophers of time have distinguished two different views about the nature of time, which have been called the A- and the B-theories of time respectively.[85] According to the A-theory, temporal moments may be classed as past, present, and future, and only that moment which is present exists. Past moments and the things or events which occupy them have passed away and no longer exist; future moments, things and events have not yet come to be and so do not yet exist. On the A-theory of time, things come into and go out of being, and thus temporal becoming is a real and objective feature of reality.

By contrast, on the B-theory of time the distinction between past, present and future is a subjective illusion of human consciousness. All things or events in time are equally real and existent, and moments, things and events merely stand to one another in tenseless relations of *earlier than, simultaneous with,* or *later than*. Nothing ever comes into or goes out of being, and temporal becoming is an illusion.

Now all instances of causal influence over the past – whether we are talking about closed time-like curves, time travel, retro-causation, tachyonic anti-telephones, or whatever – presuppose the truth of the B-theory of time.[86] For clearly on the A-theory of time, at the time at which the effect is present, the cause is future and therefore literally non-existent. Thus, the effect just comes into being from nothing. Not only does this scenario seem absurd, but it also reduces to the first horn of Davies' dilemma with respect to the origin of the universe. The universe just came uncaused from nothing.

Thus the Gott-Li-Xin hypothesis presupposes the B-theory of time. But if one presupposes such a view of time, then Gott and Li-Xin's hypothesis becomes superfluous. For on a B-theory of time the universe never truly comes into being at all.[87] The whole four-dimensional space-time manifold just exists tenselessly, and the universe has a beginning only in the sense that a meter-stick has a beginning prior to the first centimeter. Although the space-time manifold is intrinsically temporal in that one of its four dimensions is time, nonetheless it is extrinsically timeless, in that it does not exist in an embedding hyper-time but exists tenselessly, neither coming into nor going out of being. The four-dimensional space-time manifold is in this latter sense eternal. Thus, there is no need for the device of causal loops or closed time-like

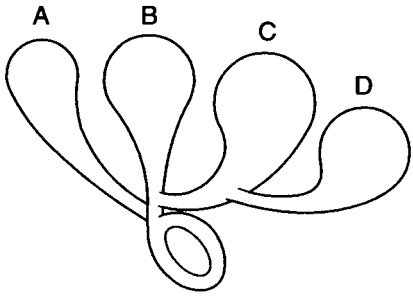

Figure 9.9 A self-creating universe. Four inflating baby universes are shown. Universes A and D have not created any baby universes. Universe C has created universe D. Universe B has created three universes: A, C and itself, B. The torus-shaped region at the bottom is a region of closed time-like curves. Such a universe neither arose from a singularity nor tunneled from nothing, but it created itself.

curves at the beginning to explain how it came into being.

Now space does not permit me to review the arguments for and against the A- and B-theories of time. I have explored this fascinating debate for the last dozen years and report my findings elsewhere.[88] Here I can only outline my reasons for affirming an A-theory of time as the most plausible view of the matter:

I. Arguments for the A-theory
   A. Linguistic tense, which is ineliminable and irreducible, mirrors the tensed facts which are characteristic of reality.[89]
   B. The experience of temporal becoming, like our experience of the external world, should be regarded as veridical.[90]

II. Refutation of arguments against the A-theory
   A. McTaggart's Paradox is based upon the illicit assumption that there should exist a unique tenseless description of reality, as well as the illicit conflation of A-theoretic becoming with a B-theoretic ontology.[91]
   B. The passage of time is not a myth, but a metaphor for the objectivity of temporal becoming, a notion which can be consistently explicated on a presentist metaphysic.[92]

III. Refutation of arguments for the B-theory
   A. Temporal becoming is wholly compatible with the mathematical core of Relativity Theory, even if its affirmation requires a different physical interpretation than the received view.[93]
   B. Time, as it plays a role in physics, is an abstraction of a richer metaphysical

reality, omitting indexical elements such as the "here" and the "now" in the interest of universalizing the formulations of natural laws.[94]

IV. Arguments against the B-theory

A. In the absence of objective distinctions between past, present and future, the relations ordering events on the B-theory are only gratuitously regarded as genuinely temporal relations of *earlier/later than*.[95]

B. The subjective illusion of temporal becoming involves itself an objective temporal becoming of contents of consciousness.[96]

C. The B-theory entails perdurantism, the view that objects have spatio-temporal parts, a doctrine which is metaphysically counter-intuitive, incompatible with moral accountability, and entails the bizarre counterpart theory of transworld identity.[97]

Given the truth of the A-theory of time, the idea that the universe is self-created, that is to say, brought itself into being via closed time-like curves, is either metaphysically impossible or else reduces to the notion that the universe sprang into existence uncaused out of nothing. Thus, I think that we are stuck with Davies' dilemma: the beginning of the universe is either an event without a cause or it is the result of a supernatural agency.

The supernaturalist alternative

Suppose we go the route of postulating some causal agency beyond space and time as being responsible for the origin of the universe. A conceptual analysis of what properties must be possessed by such an ultra-mundane cause enables us to recover a striking number of the traditional divine attributes. For as the cause of space and time, this entity must transcend space and time and therefore exist atemporally and non-spatially, at least sans the universe. This transcendent cause must therefore be changeless and immaterial, since timelessness entails changelessness, and changelessness implies immateriality. Such a cause must be beginningless and uncaused, at least in the sense of lacking any antecedent causal conditions. Ockham's Razor will shave away further causes, since we should not multiply causes beyond necessity. This entity must be unimaginably powerful, since it created the universe out of nothing.

Finally, and most strikingly, such a transcendent cause is plausibly to be regarded as personal. As Swinburne points out, there are two types of causal explanation: scientific explanations in terms of laws and initial conditions and personal explanations in terms of agents and their volitions.[98] A first state of the universe cannot have a scientific explanation, since there is nothing before it, and therefore it can be accounted for only in terms of a personal explanation. Moreover, the personhood of the cause of the universe is implied by its timelessness and

immateriality, since the only entities we know of which can possess such properties are either minds or abstract objects, and abstract objects do not stand in causal relations. Therefore the transcendent cause of the origin of the universe must be of the order of mind. This same conclusion is also implied by the origin of a temporal effect from a timeless cause. For if the cause of the universe were an impersonal set of necessary and sufficient conditions, it could not exist without its effect. The only way for the cause to be timeless and changeless but for its effect to originate *de novo* a finite time ago is for the cause to be a personal agent who freely chooses to bring about an effect without antecedent determining conditions. Thus, we are brought, not merely to a transcendent cause of the universe, but to its personal creator.

Naturalistic objections

The naturalist, of course, will be reluctant to take on board such metaphysical baggage. But think of the alternative: that the universe came into being uncaused out of nothing. That seems metaphysically absurd. The naturalist philosopher of science Bernulf Kanitscheider remonstrates, "If taken seriously, the initial singularity is in head-on collision with the most successful ontological commitment that was a guiding line of research since Epicurus and Lucretius," namely, *out of nothing nothing comes*, which Kanitscheider calls "a metaphysical hypothesis which has proved so fruitful in every corner of science that we are surely well-advised to try as hard as we can to eschew processes of absolute origin."[99] Mario Bunge thinks that an absolute origin of the universe "would be unscientific, for science abides by the principles that nothing comes out of nothing or turns into nothingness . . . and that everything happens according to law rather than miracles."[100] On the basis of the first principle, Bunge, like Kantischeider, rejects the view that the universe came into being uncaused out of nothing. On the basis of the second principle, he thinks to reject theism. But while the principle that *out of nothing nothing comes* is a first principle of metaphysics as well as science, there is no incompatibility between being a theist metaphysically and a methodological naturalist scientifically; moreover, even methodological naturalism is far from unchallengeable.[101] It is difficult to see how any sensible person, particularly the naturalist, can think that the universe just sprang into existence uncaused out of nothing.

It has therefore been remarkable to observe in recent years the number of naturalists who, under the force of the evidence for an absolute beginning of the universe, have embraced the view that the universe is a surd contingent, something which popped into existence uncaused out of nothing. Quentin Smith declares, "The fact of the matter is that the most reasonable belief is that we came from nothing, by nothing and for nothing."[102] Rather than posit a cause of the origin of the universe, Smith advises, "We should instead acknowledge our foundation in nothingness and feel awe at the marvelous fact that we have a chance to participate briefly in this

incredible sunburst that interrupts without reason the reign of non-being."[103]

Sometimes attempts are made to render this remarkable hypothesis more plausible, but these are usually not very impressive. Consider, for example, Peter Atkins's account of the origin of the universe:

> Now we go back in time beyond the moment of creation, to when there was no time, and to where there was no space. . . . In the beginning there was nothing. . . . By chance there was a fluctuation, and a set of points, emerging from nothing . . . defined a time. . . . From absolute nothing, absolutely without intervention, there came into being rudimentary existence. . . . Yet the line of time collapsed, and the incipient universe evaporated, for time alone is not rich enough for existence. Time and space emerged elsewhere, but they too crumbled back into their own dust, the coalescence of opposites, or simply nothing. Patterns emerged again, and again, and again. Each time the pattern formed a time, and through their patterning into time, the points induced their own existence. . . . Sometimes chance patterned points into a space as well as a time. . . . Then, by chance, there came about our fluctuation. Points came into existence by constituting time but, this time, in this pattern time was accompanied by three dimensions of space . . . with them comes stability, later elements, and still later elephants.[104]

This account is so obviously incoherent in postulating time before time and so confused in its reification of mathematical entities that we may rightly dismiss it as the pseudo-scientific drivel that it is.[105]

Or again, when John Gribbin asserts that the origin of the universe from nothing presents no problem, since the positive energy associated with mass is precisely offset by the negative energy associated with gravitation, so that in the case of the origin of the universe we got "Not something for nothing, after all, but *nothing* for nothing,"[106] he commits himself to the absurd position that nothing exists (not even he himself!). At the very best, the fact that the universe contains counter-balancing amounts of positive and negative energy could show that the universe need not have a material cause; but it does nothing to obviate the need for an efficient cause. As Isham puts it, there is still the "need for ontic seeding" to produce the positive and negative energy, even if on balance it is naught.[107] That is why the quantum vacuum was needed as a substratum in cosmogonic theories postulating such a process.

More often naturalistic thinkers have sought to commend their view either by attacking the causal principle *whatever begins to exist has a cause* or else by arguing for the implausibility or incoherence of the existence of a cause of the universe. Attacks on the causal principle are usually based on an appeal to quantum indeterminacy. For example, virtual particles are sometimes said to constitute a counter-example to the

principle because they spring uncaused from the quantum mechanical vacuum. Wholly apart from the disputed question as to whether virtual particles really exist at all,[108] the central point to be understood here is that the quantum vacuum on which they depend for their existence is not nothing. It is for that reason that the statements frequently made with respect to Vacuum Fluctuation models that "the universe quantum tunneled into being out of nothing," or that "nothingness is unstable" to fluctuations which grow into universes, or that "the universe is a free lunch" because in this case "we got something for nothing" cannot be taken seriously, for they treat nothing as though it were something, a sort of substance possessing properties and governed by the laws of quantum physics. In fact such statements turn out to be just rhetorical flourishes which no informed scientist takes literally. The quantum vacuum, which underlies all of space-time reality, is a fluctuating sea of energy. Because the vacuum is a physical entity existing in space and time, Vacuum Fluctuation models did not envision a genuine origin of the universe out of nothing, as Kanitscheider emphasizes:

> The violent microstructure of the vacuum has been used in attempts to explain the origin of the universe as a long-lived vacuum fluctuation. But some authors have connected with this legitimate speculations [*sic*] far-reaching metaphysical claims, or at most they couched their mathematics in a highly misleading language, when they maintained 'the creation of the universe out of nothing' ....
>
> From the philosophical point of view it is essential to note that the foregoing is far from being a spontaneous generation of everything from naught, but the origin of that embryonic bubble is really a causal process leading from a primordial substratum with a rich physical structure to a materialized substratum of the vacuum. Admittedly this process is not deterministic, it includes that weak kind of causal dependence peculiar to every quantum mechanical process.[109]

Thus, quantum physics does not serve to rebut the principle that whatever begins to exist has a cause.

It is not surprising that naturalists should attack the notion of a cause of the universe, since they reject supra-natural realities independently of their motivation to justify an uncaused origin of the universe from nothing. Sometimes these critiques may be easily dismissed. For example, metaphysician John Post obviously begs the question when he claims that there cannot be a cause of the origin of the universe, since "by definition the universe contains everything there is or ever was or will be."[110] Again it is an obvious *non-sequitur* when he infers that because "the singularity cannot be caused by some earlier *natural* event or process," therefore "the universe has an uncaused beginning" and "it seems contemporary physical cosmology cannot be cited in support of the idea of a *divine* cause or creator of the universe."[111]

On the other hand, Smith realizes that the metaphysician must take seriously the

"more difficult question" of "whether or not the singularity or the Big Bang probably is an effect of a supernatural cause, God."[112] What problems, then, are there with a supernaturalist perspective? Adolf Grünbaum has argued vigorously against what he styles "the New Creation Argument" for a supernatural cause of the origin of the universe.[113] His basic *Ansatz* is based on the assumption that causal priority implies temporal priority. Since there were no instants of time prior to the Big Bang, it follows that the Big Bang cannot have a cause.[114]

It seems to me that the supernaturalist has a number of options for dealing with this objection, one of which is to hold that the transcendent cause of the universe is causally, but not temporally, prior to the Big Bang event, such that His act of causing the universe to begin to exist is simultaneous, or co-incident, with its beginning to exist. Grünbaum provides no justification for his assumption that causal priority implies temporal priority. Discussions of causal directionality deal routinely with cases in which cause and effect are simultaneous. A supernaturalist could hold that the Creator sans the universe exists changelessly and, hence, timelessly and at the Big Bang singularity created the universe along with time and space. For the Creator sans the universe, there simply is no time because there are no events of any sort; time begins with the first event, at the moment of creation.

The time of the first event would be not only the first time at which the universe exists, but also, technically, the first time at which God exists, since sans the universe God exists timelessly.[115] The moment of creation is, as it were, the moment at which God enters time. His act of creation is thus simultaneous with the origination of the universe.

In response to this suggestion, Grünbaum has opposed the following argument:[116]

1. The proponent of simultaneous, asymmetric causation must furnish a generally accepted criterion for distinguishing one of two causally connected simultaneous events as the cause of the other, if simultaneous, asymmetric causation is possible.
2. There is no generally accepted account of causal directionality.
3. Therefore, there can be no simultaneous, asymmetric cause of the Big Bang.

The argument, if successful, would eliminate all purported instances of simultaneous, asymmetric causation, not just a cause of the Big Bang.

The argument, however, is, I think, unsound because (1) is so obviously false. Why must the proponent of simultaneous, asymmetric causation furnish a *generally accepted* criterion of causal directionality in order for such causation to be possible? Is this not an extravagant demand? Grünbaum fails to appreciate that there is no generally accepted account of the direction of causation *überhaupt*, including accounts which appeal to temporal priority as a condition of causal priority. Indeed, I should dare to

say that there is no generally accepted account of causation at all today. But should we therefore infer that causation is impossible or non-existent? Compare the situation in contemporary epistemology. There is today no generally accepted account of justification or rational warrant with respect to beliefs we hold to be true; but should we therefore infer that knowledge is impossible? Deconstructionists and other post-modernists may think so, but I doubt that Grünbaum would be ready to follow in their train. There is no reason to think that the possibility of simultaneous causation depends upon our being able to come up with an uncontroversial criterion of causal directionality. Indeed, what reason is there to think that the possibility of simultaneous, asymmetric causation depends upon my being able to come up with any kind of criterion of causal directionality at all? My enunciation of a criterion for distinguishing a cause from its effect is an epistemic affair; the existence of simultaneous causation is a matter of ontology. A criterion helps us to *discern* simultaneous, asymmetric causes in the world; but to suggest that said criterion somehow *constitutes* such causal relations in reality is verificationism at its most implausible. Grünbaum has not suggested any incoherence or difficulty in simultaneous, asymmetric causation; if there are such causes in the world, they do not have to wait around for us to discover some criterion for distinguishing them. Furthermore, there is no reason to think that in order for specific cases of simultaneous, asymmetric causation to be possible or discernible, one must be able to furnish a general criterion broad enough to cover all such alleged cases. All one needs is a way of distinguishing cause from effect in the specific case. Now in the case of the hypothesis of theological creationism, we have a logically airtight means of distinguishing cause from effect, namely, it is *metaphysically impossible* for God to be caused by the world, since if God exists, His nature is such that He exists necessarily, whereas the world's existence is metaphysically contingent (as is evident from its beginning to exist). That entails that there is *no possible world* in which God is caused by the Big Bang. Hence, it is easy for the theist to explain in what sense God is causally prior to the universe or the Big Bang: God and the universe are causally related, and if the universe were not to exist, God would nevertheless exist, whereas there is no possible world in which the universe exists without God. Thus, it seems to me that Grünbaum's objection to a supernatural cause of the origin of the universe is unsuccessful.

The naturalist will perhaps raise a metaphysical objection to the scenario I have sketched of the Creator's status sans the universe. For it requires that we conceive of a timeless, personal agent, and some philosophers have argued that such a notion is self-contradictory.[117] For it is a necessary condition of personhood that an individual be capable of remembering, anticipating, reflecting, deliberating, deciding, and so forth. But these are inherently temporal activities. Therefore, there can be no atemporal persons.

The fallacy of this reasoning is that it conflates *common* properties of persons with

essential properties of persons. The sorts of activities delineated above are certainly common properties of temporal persons. But that does not imply that such properties are essential to personhood. Arguably, what is necessary and sufficient for personhood is self-consciousness and free volition, and these are not inherently temporal. In his study of divine timelessness, John Yates writes,

> The classical theist may immediately grant that concepts such as reflection, memory, and anticipation could not apply to a timeless being (nor to any omniscient being), but this is not to admit that the key concepts of consciousness and knowledge are inapplicable to such a deity . . . there does not seem to be any essential temporal element in words like . . . 'understand,' to 'be aware,' to 'know,' and so on . . . an atemporal deity could possess maximal understanding, awareness, and knowledge in a single, all-embracing vision of himself and the sum of reality.[118]

Similarly, God could possess a free, changeless intention of the will to create a universe with a temporal beginning. Thus, neither self-consciousness nor free volition entail temporality. But since these are plausibly sufficient for personhood, there is no incoherence in the notion of a timeless, personal Creator of the universe.

More recently Smith has argued that "the thesis that the universe has an originating divine cause is logically inconsistent with all extant definitions of causality and with a logical requirement upon these and all possible valid definitions or theories of causality."[119] Smith shows that the typical analyses of the causal relation in terms of temporal priority, spatio-temporal contiguity, and nomological relatedness are inapplicable to the event of God's willing that the Big Bang occur and the event of the occurrence of the Big Bang. Therefore, these two events cannot, on the customary analyses, be regarded as cause and effect. Counterfactual analyses of causation, such as David Lewis's, according to which *c* causes *e* iff (i) *c* and *e* are both events which occur and (ii) if *c* had not occurred, *e* would not have occurred, fare no better in Smith's view. For if *c* is God's willing and *e* is the Big Bang, it is true that if *e* had not occurred, then *c* would not have occurred. But this implies that the Big Bang is the cause of God's willing, which is false. Lewis avoids the problem of spurious reverse causal dependence by stipulating that if *e* had not occurred, then *c* would have occurred but failed to cause *e*. But since God is omnipotent and His willing necessarily effective, such a stipulation cannot be made in the present case. Thus, under no extant analysis of causality can God be said to cause the Big Bang.

Smith's argument may be formulated as follows:

4. If the claim that God caused the Big Bang cannot be analyzed in terms of extant definitions of causality, then God cannot have caused the Big Bang.
5. The claim that God caused the Big Bang cannot be analyzed in terms of extant

definitions of causality.

6. Therefore, God cannot have caused the Big Bang.

Is this argument sound and persuasive? I think not.

Consider premiss (4). I see no reason to think that this premiss is true. In general, arguments to the effect that some intuitively intelligible notion cannot be analyzed in terms of current philosophical theories ought to make us suspect the adequacy of those theories rather than reject the commonsense notion. The idea that God caused the universe is intuitively intelligible. A cause is, loosely speaking, something which produces something else and in terms of which the thing that is produced can be explained. That notion certainly applies to God's causing the universe. Indeed, God's causing certain effects is analogous to our acting as agents to bring about certain effects. We certainly conceive of ourselves as causes, and, intuitively, God should count as a cause as well. But Smith's argument, if successful, could be generalized to prove that God is not a cause of anything whatsoever. If God's acting as a cause cannot be analyzed in terms of current philosophical definitions of causation, then so much the worse for those definitions! That only shows that the definitions need to be revised. Indeed, the standard procedure in terms of which proposed definitions of causality are assessed is to postulate counter-examples of intuitively obvious cases of causation and then show how the definition fails to accommodate these examples. In the same way, if God's being a cause cannot be accommodated by some philosophical definition of causality, then that plausibly constitutes a counter-example to the definition which shows its inadequacy as a general metaphysical analysis of the causal relation, however adequate it might be for scientific purposes.[120]

Moreover, there is no reason to believe that we have arrived at the final and correct analysis of causation. In fact, there is good reason to believe the opposite. The definitions discussed by Smith are exclusively concerned with natural, even physical, causes. They were not intended to cover such re-condite cases as divine causation of the origin of the universe. It is hardly surprising, therefore, that these analyses should fail to capture this notion.

Finally, Smith just assumes that an analysis of the causal relation can be given. But it could be held that such a relation is conceptually primitive, in which case we should not expect a successful reductive analysis to exist which will cover all cases. The plethora of competing extant analyses and the recognized deficiencies of all of them lend credibility to this viewpoint.

What about premiss (5)? It seems to me that there are analyses of causation, however inadequate, which can accommodate God's causing the Big Bang. Consider Lewis's analysis of causation. According to Lewis, *c* causes *e* if and only if *c* and *e* are both events which occur and if *c* had not occurred *e* would not have occurred. Now God's willing the Big Bang clearly satisfies this definition: God's willing and the Big Bang are both events which occur, and if God's willing had not occurred, the Big Bang would

not have occurred. But Smith rejoins, "But if the Big Bang had not occurred, God's willing would not have occurred. So is the Big Bang the cause of God's willing?" Obviously not; but what this calls into question is the *adequacy* of Lewis's analysis, not whether divine causation satisfies it. Lewis remedies the problem by stipulating that if *e* had not occurred, *c* would still have occurred but failed to cause *e,* a remedy which will not work for divine causation. Actually Lewis's remedy will not work for many natural causes either, since in some cases the counterfactual, "If *e* had not occurred, *c* would not have occurred" is true. So what Lewis's definition gives is not an analysis of "*c* causes *e*" but rather an analysis of "*c* and *e* are causally related," and it fails to specify the *direction* of causation. But the theist faces no problem there: for, as we have said, it is metaphysically impossible for God's willing to have an external cause. There is no possible world in which the Big Bang causes God's volition. Therefore, given Lewis's analysis of "*c* and *e* are causally related" and the impossibility of the Big Bang's causing God's willing, it follows that God's willing causes the Big Bang. Thus, divine causation satisfies Lewis's definition of causality.

Again, there are analyses of agent causation which are even more relevant in the case of divine causation than the analyses surveyed by Smith. Smith considers exclusively event causation, but it may be disputed whether this is the correct conception to apply to God's case. Smith contends that considerations of agent causation are not germane to the discussion because we are not concerned with the relation between God (the agent) and His act of willing (the effect), but with the relation between His act of willing (an event) and the Big Bang (an event). But not all proponents of agent causation construe that notion as a relation between an agent and his volitions. Some proponents of agent causation hold that an agent does not cause his volitions, but that by freely willing he brings about some intended event. In the case at hand, God brings about the Big Bang by an exercise of His active power. Thus, it is simply wrong-headed to think of the Big Bang as caused by the event of God's willing rather than by God Himself.[121]

Thus, neither (4) nor (5) commends itself to us as more plausibly true than its contradictory. Smith recognizes these deficiencies of his argument, but he falls back to what he considers an impregnable position: "*c* is a cause of *e*" entails "*c* is not a logically sufficient condition of *e*."[122] This entailment precludes God's being the cause of the Big Bang because God's willing that the Big Bang occur is a logically sufficient condition of the Big Bang. This is because God is omnipotent, and thus necessarily His will is successful. There is no possible world in which God wills the Big Bang and yet the Big Bang fails to occur. Therefore, God cannot be the cause of the Big Bang.

This argument seems quite fanciful. If successful, it can be generalized to show that God cannot cause anything. Thus, precisely *because* He is omnipotent, God is utterly impotent – a strange inference! If being omnipotent entails inability to cause anything, then we are using "cause" in a highly technical sense which is not

incompatible with God's bringing about the Big Bang, which is, after all, the issue. Whether or not God "causes" the Big Bang, it is still up to Him whether it occurs or not, and it occurs only in virtue of His willing that it occur. If it seems that bringing about the Big Bang does involve a causal relation, then we shall simply reject Smith's entailment principle. Only someone who is already a naturalist would be tempted to think that that principle is true. Thus, Smith's argument is either questionbegging or not incompatible with God's bringing about the Big Bang.

Smith considers such a response and insists that it is the theist who begs the question, since in every other case of causation causes are not logically sufficient conditions of their effects. There is, he says, no justification for exempting God's alleged acts of causation from this principle. We need to have some independent reason for thinking that the relation between God and the Big Bang is a causal relation. Three things may be said about this response: First, since only God is omnipotent, it is hardly surprising that His case should be the sole exception to the principle that causes are not logically sufficient for their effects. God is so exceptional a being that He will in general not fit into our customary schemata. For example, it is a general principle that "S believes *p*" is not a logically sufficient condition of "*p*." But since God is essentially omniscient, in God's case His believing p is a logically sufficient condition of *p*. Should we therefore conclude that God has no beliefs? In the same way, because God is omnipotent, are we to think that His will has no effects? Second, there are other plausible counter-examples to Smith's principle. For example, change is plausibly a cause of the existence of time, at least on a relational view of time. The occurrence of events actually brings time into existence. If there were an absolutely quiescent state then time would not exist. But if a change occurs, time is immediately produced. Such a relation is plausibly causal; it is certainly not like the purely logical relation between, say, a two-dimensional figure's having three sides and its having three angles. Time is something altogether distinct from change, since time can go on, even most relationalists agree, even though change should cease.[123] Thus, change, should it occur, would seem to cause time to exist. Yet change necessarily causes time: there is no possible world in which change is going on without time. Change is thus logically sufficient for the existence of time, but is also plausibly a cause of time's existence. Third, the reason that the relation between God and the Big Bang – or any other event He brings about – is causal is the close resemblance between God and ourselves as agents. Doubtless our deepest intuitions about causality are rooted in our own ability to bring about effects by an intentional exertion of our power. But God is a personal agent like ourselves. The difference between Him and us is that His power is so great that He is infallible in bringing about His undertakings. Is His status as a cause now to be doubted because He is infallible? Hardly! In short I do not think that Smith's objection poses a serious obstacle to thinking that the Big Bang has a supernatural or divine cause.

All of the above objections have been considered as attempted justification of the

apparently incredible position that the universe sprang into being uncaused out of nothing. But I, for one, find the premisses of those objections far less perspicuous than the proposition that *whatever begins to exist has a cause.* It is far more plausible to deny one of those premisses than to affirm what Hume called the "absurd Proposition" that something might arise without a cause,[124] that the universe, in this case, should pop into existence uncaused out of nothing.

Conclusion

We can summarize our argument as follows:

7. Whatever begins to exist has a cause of its existence.
8. The universe began to exist.
9. Therefore, the universe has a cause of its existence.

Premiss (7) is an intuitively grasped, metaphysical first principle. Premiss (8) is supported by the inductive evidence of contemporary cosmology and enjoys greater plausibility in light of that evidence than its contradictory. An analysis of what it is to be cause of the universe reveals that

10. If the universe has a cause of its existence, then an uncaused, personal Creator of the universe exists, who sans the universe is beginningless, changeless, immaterial, timeless, spaceless, and enormously powerful.

From (9) and (10), it follows that

11. Therefore, an uncaused, personal Creator of the universe exists, who sans the universe is beginningless, changeless, immaterial, timeless, spaceless, and enormously powerful.

And this, as Thomas Aquinas laconically remarked, is what everyone means by "God."[125]

Notes

1 Norman Malcolm, *Ludwig Wittgenstein: A Memoir* (London: Oxford University Press, 1958),70.

2 Aristotle wrote, "For it is owing to their wonder that men both now begin and at first began to philosophize; they wondered originally at the obvious difficulties, then advanced little by little and stated difficulties about the greater matters, e.g. about the phenomena of the moon and those of the sun and the stars, and about the genesis of the universe" (*Metaphysics* A. 2. 982b10–15).

3 Derek Parfit, "Why Anything? Why This?," *London Review of Books* 20/2 (January 22, 1998),

24.

4 Gottfried Wilhelm Leibniz, "The Principles of Nature and of Grace, Founded on Reason," in *The Monadology and Other Philosophical Writings*, trans. Robert Latta (London: Oxford University Press, 1951), 415; idem, "The Monadology," in *Monadology and Other Philosophical Writings*, 237–9.

5 For this analysis of so-called factual necessity, see John Hick, "God as Necessary Being," *Journal of Philosophy* 57 (1960): 733–4.

6 David Hume, *Dialogues concerning Natural Religion*, ed. with an Introduction by Norman Kemp Smith, Library of Liberal Arts (Indianapolis: Bobbs-Merrill, 1947), pt. IX, 190.

7 Ibid.

8 Bertrand Russell and F. C. Copleston, "The Existence of God," in *The Existence of God*, ed. with an Introduction by John Hick, Problems of Philosophy Series (New York: Macmillan, 1964), 175.

9 A. Einstein, "Cosmological Considerations on the General Theory of Relativity," in *The Principle of Relativity*, by A. Einstein, *et al.*, with Notes by A. Sommerfeld, trans. W. Perrett and J. B. Jefferey (rep. ed.: New York: Dover Publications, 1952), 177–88.

10 A. Friedman, "Über die Krümmung des Raumes," *Zeitschrift für Physik* 10 (1922): 377–86; G. Lemaître, "Un univers homogène de masse constante et de rayon croissant, rendant compte de la vitesse radiale des nébuleuses extragalactiques," *Annales de la Société scientifique de Bruxelles* 47 (1927): 49–59.

11 Gregory L. Naber, *Spacetime and Singularities: an Introduction* (Cambridge: Cambridge University Press, 1988), 126–7.

12 E. Hubble, "A Relation between Distance and Radial Velocity among Extragalactic Nebulae," *Proceedings of the National Academy of Sciences* 15 (1929): 168–73.

13 John A. Wheeler, "Beyond the Hole," in *Some Strangeness in the Proportion*, ed. Harry Woolf (Reading, Mass.: Addison-Wesley, 1980), 354.

14 P. C. W. Davies, "Spacetime Singularities in Cosmology and Black Hole Evaporation," in *The Study of Time III*, ed. J. T. Fraser (N.Y.: Springer-Verlag, 1978), 78–9.

15 As Gott, Gunn, Schramm, and Tinsley write,

> the universe began from a state of infinite density about one Hubble time ago. Space and time were created in that event and so was all the matter in the universe. It is not meaningful to ask what happened before the big bang; it is somewhat like asking what is north of the North Pole. Similarly, it is not sensible to ask where the big bang took place. The point-universe was not an object isolated in space; it was the entire universe, and so the only answer can be that the big bang happened everywhere (J. Richard Gott III, James E. Gunn, David N. Schramm, and Beatrice M. Tinsley, "Will the Universe Expand Forever?," *Scientific American* [March 1976], 65).

The Hubble time is the time since the singularity if the rate of expansion has been constant. The singularity is a point only in the sense that the distance between any two points in the singularity is zero. Anyone who thinks there must be a place in the universe where the Big Bang occurred still has not grasped that it is space itself which is expanding; it is the two-dimensional *surface* of an inflating balloon which is analogous to three-dimensional space. The spherical surface has no center and so no location where the expansion begins. The analogy of the North Pole with the beginning of time should not be pressed, since the North Pole is not an edge to the surface of the globe; the beginning of time is more like the apex of a cone. But the idea is that just as one cannot go further north than

the North Pole, so one cannot go earlier than the initial singularity.

16 John Barrow and Frank Tipler, *The Anthropic Cosmological Principle* (Oxford: Clarendon Press, 1986), 442.

17 Arthur Eddington, *The Expanding Universe* (New York: Macmillan, 1933), 124.

18 Ibid., 178.

19 Quentin Smith, "The Uncaused Beginning of the Universe," in *Theism, Atheism, and Big Bang Cosmology*, by William Lane Craig and Quentin Smith (Oxford: Clarendon Press, 1993), 120.

20 Hubert Reeves, Jean Audouze, William A. Fowler, and David N. Schramm, "On the Origin of Light Elements," *Astrophysical Journal* 179 (1973): 912.

21 Fred Hoyle, *Astronomy Today* (London: Heinemann, 1975), 165.

22 Fred Hoyle, *Astronomy and Cosmology: A Modern Course* (San Francisco: W. H. Freeman, 1975), 658.

23 H. Bondi and T. Gold, "The Steady State Theory of the Expanding Universe," *Monthly Notices of the Royal Astronomical Society* 108 (1948): 252–70; F. Hoyle, "A New Model for the Expanding Universe," *Monthly Notices of the Royal Astronomical Society* 108 (1948): 372–82.

24 As Jaki points out, Hoyle and his colleagues were inspired by "openly antitheological, or rather anti-Christian motivations" (Stanley L. Jaki, *Science and Creation* [Edinburgh: Scottish Academic Press, 1974], 347). Martin Rees recalls his mentor Dennis Sciama's dogged commitment to the Steady State model: "For him, as for its inventors, it had a deep philosophical appeal – the universe existed, from everlasting to everlasting, in a uniquely self-consistent state. When conflicting evidence emerged, Sciama therefore sought a loophole (even an unlikely seeming one) rather as a defense lawyer clutches at any argument to rebut the prosecution case" (Martin Rees, *Before the Beginning*, with a Foreword by Stephen Hawking [Reading, Mass.: Addison-Wesley, 1997], 41). The phrase "from everlasting to everlasting" is the Psalmist's description of God (Ps. 90.2). Rees gives a good account of the discoveries leading to the demise of the Steady State model.

25 See, e.g., E. M. Lifschitz and I. M Khalatnikov, "Investigations in Relativist Cosmology," *Advances in Physics* 12 (1963): 207.

26 As evident from the sentiments expressed by John Gribbin:

> The biggest problem with the Big Bang theory of the origin of the universe is philosophical – perhaps even theological – what was there before the bang? This problem alone was sufficient to give a great initial impetus to the Steady State theory; but with that theory now sadly in conflict with the observations, the best way round this initial difficulty is provided by a model in which the universe expands from a singularity, collapses back again, and repeats the cycle indefinitely (John Gribbin, "Oscillating Universe Bounces Back," *Nature* 259 [1976]: 15).

> Scientists not infrequently misexpress the difficulty posed by the beginning of the universe as the question of what existed before the Big Bang (which invites the easy response that there was no "before"). The real question concerns the causal conditions of this event, why the universe exists rather than nothing.

27 R. Penrose, "Gravitational Collapse and Space-Time Singularities," *Physical Review Letters* 14 (1965): 57–9; S. W. Hawking and R. Penrose, in *The Large-Scale Structure of Space-Time*, ed. S. W. Hawking and G. F. R. Ellis (Cambridge: Cambridge University Press, 1973), 266.

28 Stephen Hawking and Roger Penrose, *The Nature of Space and Time*, The Isaac Newton

Institute Series of Lectures (Princeton, N. J.: Princeton University Press, 1996), 20.

29 Alan Guth and Mark Sher, "The Impossibility of a Bouncing Universe," *Nature* 302 (1983): 505–6; Sidney A. Bludman, "Thermodynamics and the End of a Closed Universe," *Nature* 308 (1984): 319–22.

30 Associated Press News Release, 9 January 1998.

31 Ibid. See also James Glanz, "Astronomers See a Cosmic Antigravity Force at Work," *Science* 279 (February 27, 1998): 1298–9.

32 I. D. Novikov and Ya. B. Zeldovich, "Physical Processes near Cosmological Singularities," *Annual Review of Astronomy and Astrophysics* 11 (1973): 401–2.

33 Joseph Silk, *The Big Bang*, 2nd ed. (San Francisco: W. H. Freeman, 1989), 311–12.

34 See D. Hochberg, C. Molina-Paris, and M. Visser, "Tolman Wormholes Violate the Strong Energy Condition," *Physical Review D* 59 (1999): forthcoming.

35 One thinks, for example, of the late Carl Sagan on his *Cosmos* television series propounding this model and reading from Hindu scriptures about cyclical Brahman years in order to illustrate the oscillating universe, but with nary a hint to his viewers about the difficulties attending this model.

36 Christopher Isham, "Creation of the Universe as a Quantum Process," in *Physics, Philosophy and Theology: a Common Quest for Understanding*, ed. R. J. Russell, W. R. Stoeger and G. V. Coyne (Vatican City: Vatican Observatory, 1988), 378.

37 Edward Tryon, "Is the Universe a Vacuum Fluctuation?" *Nature* 246 (1973): 396–7.

38 A. Guth, "Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems," *Physical Review D* 23 (1981): 247–56.

39 See Isham, "Creation of the Universe," 385–7.

40 J. R. Gott III, "Creation of Open Universes from de Sitter Space," *Nature* 295 (1982): 304–7.

41 Christopher Isham, "Space, Time, and Quantum Cosmology," paper presented at the conference "God, Time, and Modern Physics," March 1990.

42 Christopher Isham, "Quantum Cosmology and the Origin of the Universe," lecture presented at the conference "Cosmos and Creation," Cambridge University, 14 July 1994.

43 See, e.g., A. D. Linde, "The Inflationary Universe," *Reports on Progress in Physics* 47 (1984): 925–86; idem, "Chaotic Inflation," *Physics Letters* 1298 (1983): 177–81.

44 Robert Brandenberger, personal communication.

45 Linde, "Inflationary Universe," 976.

46 A. Borde and A. Vilenkin, "Eternal Inflation and the Initial Singularity," *Physical Review Letters* 72 (1994): 3305, 3307.

47 Andrei Linde, Dmitri Linde, and Arthur Mezhlumian, "From the Big Bang Theory to the Theory of a Stationary Universe," *Physical Review D* 49 (1994): 1783–826.

48 Borde and Vilenkin, "Eternal Inflation," 3307.

49 A. Vilenkin, "Birth of Inflationary Universes," *Physical Review D* 27 (1983): 2854. See J. Hartle and S. Hawking, "Wave Function of the Universe," *Physical Review D* 28 (1983): 2960–75; A. Vilenkin, "Creation of the Universe from Nothing," *Physical Letters* 117B (1982): 25–8.

50 Stephen Hawking, *A Brief History of Time* (New York: Bantam Books, 1988), 9.

51 Ibid., 46.

52 Ibid., 136.

53 Ibid., 140–1.

54 John D. Barrow, *Theories of Everything* (Oxford: Clarendon Press, 1991), 68.

55 Ibid., 67–8.

56 Hartle and Hawking, "Wave Function of the Universe," 2961; Hawking and Penrose, *Nature of Space and Time*, 85.

57 Vilenkin, "Creation of the Universe," 26.

58 Hawking and Penrose, *Nature of Space and Time*, 20.

59 See my "Hartle–Hawking Cosmology and Atheism," *Analysis* 57 (1997): 291–5. With respect to determining the wave function of the universe Bryce DeWitt says, "Here the physicist must play God" (B. DeWitt, "Quantum Gravity," *Scientific American* 249 [1983]: 120).

60 Vilenkin, "Birth of Inflationary Universes," 2851.

61 Hawking, *Brief History of Time*, 136.

62 Barrow, *Theories of Everything*, 66–7.

63 Hawking, *Brief History of Time*, 138–9.

64 Hawking and Penrose, *Nature of Space and Time*, 121.

65 Ibid., 3–4. Cf. his comment, "I . . . am a positivist who believes that physical theories are just mathematical models we construct, and that it is meaningless to ask if they correspond to reality, just whether they predict observations" (Stephen Hawking, "The Objections of an Unashamed Positivist," in *The Large, the Small, and the Human*, by Roger Penrose [Cambridge: Cambridge University Press, 1997], 169).

66 Hawking and Penrose, *Nature of Space and Time*, 121; cf. 4.

67 Ibid., 53–5.

68 As pointed out by Christopher Isham, "Quantum Theories of the Creation of the Universe," in *Quantum Cosmology and the Laws of Nature*, ed. R. J. Russell, N. Murphey, and C. J. Isham (Vatican City: Vatican Observatory, 1993), 56.

69 John D. Barrow, *The Origin of the Universe* (New York: Basic Books, Harper Collins, 1994), 113.

70 Roger Penrose, "Some Remarks on Gravity and Quantum Mechanics," in *Quantum Structure of Space and Time*, ed. M. J. Duff and C. J. Isham (Cambridge: Cambridge University Press, 1982), 4.

71 Ibid., 5.

72 Weyl curvature is the curvature of space-time which is not due to the presence of matter and is described by the Weyl tensor. Space-time curvature due to matter is described by the Ricci tensor. Together they make up the Riemann tensor giving the metric for space-time.

73 Hawking and Penrose, *Nature of Space and Time*, 129.

74 Ibid., 130.

75 Penrose, "Remarks on Gravity," 5.

76 Ibid.

77 Roger Penrose, "Time-Asymmetry and Quantum Gravity," in *Quantum Gravity 2*, ed. C. J. Isham, R. Penrose, and D. W. Sciama (Oxford: Clarendon Press, 1981), 249; cf. Hawking and Penrose, *Nature of Space and Time*, 34–5.

78 Penrose, "Time-Asymmetry," 249.

79 Some recent efforts have been made to describe a pre-Big Bang universe in terms of superstring or M-theory (M. Gasperini, "Looking Back in Time beyond the Big Bang," *Modern Physics Letters A* [forthcoming]; idem, "Inflation and Initial Conditions in the Pre-Big Bang Scenario," [pre-print] on the assumption of a "duality-symmetry" which associates a geometric mirror image with the familiar post–Big Bang expanding space-time geometry. Apart from the problems that there seems to be no way to join the pre- and post-Big Bang

eras together nor any way to smooth the transition to a matter-dominated universe, the scenario is based on a non-existent theory and so cannot even begin to be a plausible alternative.

80 Paul Davies, "The Birth of the Cosmos," in *God, Cosmos, Nature and Creativity*, ed. Jill Gready (Edinburgh: Scottish Academic Press, 1995), 8–9.

81 J. Richard Gott III and Li-Xin Li, "Can the Universe Create Itself?" *Physical Review D* 58 (2) (1998): 023501–1.

82 Ibid.

83 Ibid.

84 Ibid.

85 For a helpful introduction to these two competing perspectives, see Richard M. Gale, "The Static versus the Dynamic Temporal: Introduction," in *The Philosophy of Time: A Collection of Essays*, ed. Richard M. Gale (New Jersey: Humanities Press, 1968), 65–85.

86 See discussion in William Lane Craig, *Divine Foreknowledge and Human Freedom: The Coherence of Theism I: Omniscience*, Studies in Intellectual History 19 (Leiden: E. J. Brill, 1990), 150–6.

87 This is the salient point of Grünbaum's critique of the inference to a First Cause of the origin of the universe (Adolf Grünbaum, "A New Critique of Theological Interpretations of Physical Cosmology," *British Journal for the Philosophy of Science* [forthcoming]). As a B-theorist Grünbaum does not believe that the universe ever came into being, even if it had a first temporal interval. As he elsewhere writes, "coming *into* being (or 'becoming') is *not* a property of *physical* events themselves but only of human or conscious awareness of these events" (idem, "The Anisotropy of Time," in *The Nature of Time*, ed. T. Gold [Ithaca, N.Y.: Cornell University Press, 1967], 153). What Grünbaum fails to see, however, is that the claim that an absolute beginning of the universe entails that the universe came into being is rooted, not in the presupposition of the so-called Spontaneity of Nothingness, but in an A-theory of time.

88 See my companion volumes *The Tensed Theory of Time: a Critical Examination*, Synthese Library (Dordrecht: Kluwer Academic Publishers, forthcoming); *The Tenseless Theory of Time: a Critical Examination*, Synthese Library (Dordrecht: Kluwer Academic Publishers, forthcoming).

89 For an outstanding defense of this point, see Quentin Smith, *Language and Time* (New York: Oxford University Press, 1991). See also my "Tense and the New B-Theory of Language," *Philosophy* 71 (1996): 5–26; "The New B-Theory's *Tu Quoque* Argument," *Synthese* 107 (1996): 249–69; and "On Truth Conditions of Tensed Sentence Types," *Synthese* (forthcoming).

90 One of the most eloquent spokesmen for this point of view has been George Schlesinger, *Aspects of Time* (Indianapolis: Hackett, 1980), 34–9, 138–9. See also my "Tensed Time and our Differential Experience of the Past and Future," *Southern Journal of Philosophy* (forthcoming); "The Presentness of Experience," in *Time, Creation, and World Order*, ed. Mogens Wegener (Aarhus, Denmark: University of Aarhus Press, forthcoming); and "On Wishing It Were Now Some Other Time," *Philosophy and Phenomenological Research* (forthcoming).

91 The most helpful here are still C. D. Broad, *An Examination of McTaggart's Philosophy*, 2 vols. (Cambridge: Cambridge University Press, 1938) and Michael Dummett, "A Defense of McTaggart's Proof of the Unreality of Time," *Philosophical Review* 69 (1960): 497–504. See also my "McTaggart's Paradox and the Problem of Temporary Intrinsics," *Analysis* 58

(1998): 122–7.

92 This point needs further work, but see Arthur N. Prior, "Changes in Events and Changes in Things," in *Papers on Time and Tense* (Oxford: Clarendon Press, 1968), 1–14, and Andros Loizou, *The Reality of Time* (Brookfield, Ver.: Gower, 1986), 44–5. See also my "In Defense of Presentism," in *Philosophy of Time*, ed. L. Nathan Oaklander and Quentin Smith (forthcoming); "The Extent of the Present," *International Studies in the Philosophy of Science* (forthcoming).

93 See Smith, *Language and Time*, chap. 7. See also my "Divine Eternity and the Special Theory of Relativity," in *God and Time*, ed. Gregory Ganssle and David Woodruff (New York: Oxford University Press, forthcoming).

94 See remarks of Max Black, Review of *The Natural Philosophy of Time, Scientific American* 206 (April, 1962), 181–3.

95 For adumbrations of this argument see Richard Gale, *The Language of Time*, International Library of Philosophy and Scientific Method (London: Routledge and Kegan Paul, 1968), 90–7 and D. H. Mellor, *Real Time* (Cambridge: Cambridge University Press, 1981) 140. For a fuller development see my "Tooley on Tense and Temporal Priority," *Analysis* (forthcoming).

96 Again, this point needs to be better developed, but see Peter Geach, "Some Problems about Time," in *Logic Matters* (Berkeley: University of California Press, 1972), 306, and James McGilvray, "A Defense of Physical Becoming," *Erkenntnis* 14 (1979): 275–99.

97 See the excellent study by Trenton Merricks, "Endurance and Indiscernibility," *Journal of Philosophy* 91 (1994): 165–84; see further Mark Hinchliff, "The Puzzle of Change," paper presented at the Pacific Division meeting of the American Philosophical Association, April 2, 1994; Delmas Lewis, "Persons, Morality, and Tenselessness," *Philosophy and Phenomenological Research* 47 (1986): 305–9; Peter Van Inwagen, "Four Dimensional Objects," *Noûs* 24 (1990): 245–55.

98 Richard Swinburne, *The Existence of God*, rev. ed. (Oxford: Clarendon Press, 1991), 32–48.

99 Bernulf Kanitscheider, "Does Physical Cosmology Transcend the Limits of Naturalistic Reasoning?" in *Studies on Mario Bunge's "Treatise,"* ed. P. Weingartner and G. J. W. Doen (Amsterdam: Rodopi, 1990), 344.

100 Mario Bunge, *Treatise on Basic Philosophy*, vol. 7 :*Epistemology and Methodology III: Philosophy of Science and Technology:* Part I: *Formal and Physical Sciences* (Dordrecht: D. Reidel, 1985), 238–9.

101 See the very interesting recent discussions about the warrant for methodological naturalism in science, e.g., Paul de Vries, "Naturalism in the Natural Sciences: A Christian Perspective," *Christian Scholar's Review* 15 (1986): 388–96; Alvin Plantinga, Howard J. Van Till, Pattle Pun, and Ernan McMullin, "Symposium: Evolution and the Bible," *Christian Scholar's Review* 21 (1991): 8–109; William Hasker, "Evolution and Alvin Plantinga," *Perspectives on Science and Christian Faith* 44 (1992): 150–62; Alvin Plantinga, "On Rejecting The Theory of Common Ancestry: A Reply to Hasker," *Perspectives on Science and Christian Faith* 44 (1992): 258–63; Alvin Plantinga, "Methodological Naturalism," paper presented at the symposium "Knowing God, Christ, and Nature in the Post-Positivistic Era," University of Notre Dame, April 14–17, 1993; J. P. Moreland, "Theistic Science and Methodological Naturalism," in *The Creation Hypothesis*, ed. J. P. Moreland (Downer's Grove, Ill.: Inter-Varsity Press, 1994), 41–66; J. P. Moreland, Stephen C. Meyer, and Richard H. Bube, "Conceptual Problems and the Scientific Status of Creation Science: a Discussion," *Perspectives on Science and Christian Faith*

46 (1994): 2–25.

102 Smith, "Uncaused Beginning of the Universe," 135. Elsewhere he has written,

> [This world] exists nonnecessarily, improbably, and causelessly. It exists *for absolutely no reason at all*. . . . The impact of this captivated realization upon me is overwhelming. I am completely stunned. I take a few dazed steps in the dark meadow, and fall among the flowers. I lie stupefied, whirling without comprehension in this world through numberless worlds other than this one" (Quentin Smith, *The Felt Meanings of the World* [Lafayette, Ind.: Purdue University Press, 1986], 300–1).

In *Theism, Atheism, and Big Bang Cosmology*, Smith claimed that the universe came into being uncaused out of nothing at the Planck time; but he has since recanted that position under the realization that the whole field of quantum cosmology is then studying a complete fiction!

103 Smith, "Uncaused Beginning of the Universe," 135.

104 Peter W. Atkins, *Creation Revisited* (New York: W. H. Freeman, 1992), 129, 149–51.

105 John Leslie asks incredulously, "How could such nonsense have been churned out by the author of *Physical Chemistry*, a superb textbook?" (John Leslie, "Is It All Quite Simple?," *Times Literary Supplement*, 29 January, 1993, 3). For a good critique of Atkins, see Keith Ward, *God, Chance, and Necessity* (Oxford: One World, 1996), chap. 1.

106 John Gribbin, *In Search of the Big Bang* (New York: Bantam Books, 1986), 374.

107 Isham, "Quantum Cosmology and the Origin of the Universe," 8.

108 See Robert Weingard, "Do Virtual Particles Exist?" in *Proceedings of the Philosophy of Science Association*, 2 vols., ed. Peter Asquith and Thomas Nichols (East Lansing, Mich.: Philosophy of Science Association, 1982), 1: 235–42.

109 Kanitscheider, "Physical Cosmology," 346–7.

110 John Post, *Metaphysics: a Contemporary Introduction* (New York: Paragon House, 1991), 85.

111 Ibid., 87.

112 Smith, "Uncaused Beginning of the Universe," 120.

113 Adolf Grünbaum, "The Pseudo-Problem of Creation in Physical Cosmology," *Philosophy of Science* 56 (1989): 373–94. For a response, see William Lane Craig, "The Origin and Creation of the Universe: a reply to Adolf Grünbaum," *British Journal for the Philosophy of Science* 43 (1992): 233–40.

114 Adolf Grünbaum, "Creation as a Pseudo-Explanation in Current Physical Cosmology," *Erkenntnis* 35 (1991): 233–54. For a response, see William Lane Craig, "Prof. Grünbaum on Creation," *Erkenntnis* 40 (1994): 325–41.

115 Brian Leftow puts this nicely when he writes,

> If God existed in time once time existed and time had a first moment, then God would have a first moment of existence: there would be a moment before which He did not exist, because there was no 'before' that moment. . . . Yet even if He . . . had a first moment of existence, one could still call God's existence unlimited were it understood that He would have existed even if time did not. For as long as this is true, we cannot infer from God's having had a first moment of existence that God *came into* existence or would not have existed save if time did (Brian Leftow, *Time and Eternity*, Cornell Studies in Philosophy of Religion [Ithaca, N.Y.: Cornell University Press, 1991], 269; cf. 201).

Senor has dubbed such a model of divine eternity "accidental temporalism" (Thomas D. Senor, "Divine Temporality and Creation *ex nihilo*," *Faith and Philosophy* 10 [1993]: 88). See

further William Lane Craig, "Timelessness and *Creation*," *Australasian Journal of Philosophy* 74 (1996): 646–56.

116 Adolf Grünbaum, "Some Comments on William Craig's 'Creation and Big Bang Cosmology'," *Philosophia naturalis* 31 (1994): 225–36. For a response, see William Lane Craig, "A Response to Grünbaum on Creation and Big Bang Cosmology," *Philosophia naturalis* 31 (1994): 237–49.

117 See discussion and references in William Lane Craig, "Divine Timelessness and Personhood," *International Journal for Philosophy of Religion* 43 (1998): 109–24.

118 John C. Yates, *The Timelessness of God* (Lanham, Md.: University Press of America, 1990), 173.

119 Quentin Smith, "Causation and the Logical Impossibility of a Divine Cause," *Philosophical Topics* 24 (1996): 169–70.

120 In Quentin Smith, "The Concept of a Cause of the Universe," *Canadian Journal of Philosophy* 23 (1993): 1–24, Smith actually arrives at this conclusion himself. He states,

> extant definitions of causality are incorrect since they do not cohere in the proper way with the concept of a cause of the universe . . . .
>
> This entails that either there is some other (as yet undiscovered) definition of a cause that is correct or that a cause is indefinable. In the latter case, the concept of a cause would be primitive and the causal relation a simple relation known only by ostension (as is arguably the case with such relations as *being in contact with* or *being earlier than*). I know of no means of discovering or formulating a correct definition of a cause and know of no reason to think that there is such a definition. Accordingly, I think it is reasonable to conclude that the causal relation is indefinable.
>
> One way to avoid this conclusion would be to reject the assumption that the various examples of causes of the big bang . . . are genuine examples of causes. . . . I would say that claims that God's creation of the big bang singularity and other examples given . . . are not cases of possible causation are counterintuitive and are *ad hoc* attempts to retain a counterexampled theory. It is more plausible to think that a cause cannot be defined than to think that a mind's creation of a big bang singularity could not be a causal act (Ibid., 1, 24).

Smith came to think God's relation to the Big Bang is not causal because no cause is logically sufficient for its effect. But Smith does not justify why the actions of an omnipotent being would not be exceptions to this rule.

121 See J. P. Moreland, "Libertarian Agency and the Craig/Grünbaum Debate about Theistic Explanation of the Initial Singularity," *American Catholic Philosophical Quarterly* 81 (1998): 539–54. I am indebted to my colleague for several interesting discussions pertinent to agency and creation.

122 Smith, "Logical Impossibility of a Divine Cause," 176.

123 Sydney Shoemaker, "Time without Change," *Journal of Philosophy* 66 (1969): 363–81.

124 David Hume to John Stewart, February, 1754, in *The Letters of David Hume*, 2 vols., ed. J. Y. T. Greig (Oxford: Clarendon Press, 1932), 1: 187.

125 Thomas Aquinas *Summa theologiae* 1a.2.3.

10 Naturalism and design

William Dembski

1. Introduction

Naturalism has become the default position for all serious inquiry. From the humanities to the natural sciences to biblical studies to the arts, inquiry is allowed to proceed only under the supposition of naturalism, the view that nature is self-sufficient. Naturalism is not atheism. To affirm that nature is self-sufficient is not to deny God's existence. God could, after all, have created the world to be self-sufficient. Nonetheless, in the interest of keeping human inquiry productive, naturalism insists that the existence and activity of God be bracketed. Naturalism affirms not so much that God does not exist as that God need not exist. And because God need not exist, intellectual honesty demands that we get about the serious work of life without invoking him. This is the received wisdom.

Naturalism comes in a variety of forms. Its principal form in Western culture is known as scientific naturalism. Scientific naturalism locates the self-sufficiency of nature in the undirected natural causes studied by science. Accordingly, scientific naturalism would have us understand the universe entirely in terms of such causes. In particular, since human beings are a part of the universe, who we are and what we do must ultimately be understood in naturalistic terms. This is not to deny our humanity. But it is to interpret our humanity as the consequence of brute material processes that were not consciously aiming at us. Nor is this to deny God. But it is to affirm that if God exists, he was marvelously adept at covering his tracks and giving no empirical evidence of ever interacting with the world. To be sure, it remains logically permissible for the scientific naturalist to affirm God's existence, but only by making God a superfluous rider on top a self-contained account of the world.

If naturalism is false, how could we know it? In the last thirty years, evidence from cosmology has suggested that a designing intelligence ordered the universe. With the rise of Big Bang cosmology, the discovery of anthropic coincidences, and the proliferation of fine-tuning arguments, the existence of a designing intelligence has gained plausibility. This same evidence, however, has been interpreted naturalistically. Stephen Hawking, for instance, seeks a theory of everything that omits

all contingency and thereby eliminates any need for fine-tuning.[1] Alan Guth's inflationary cosmology attempts to wash out the need for fine-tuning by giving the early universe a brief but explosive period of exponential growth.[2] John Barrow and Frank Tipler show that by positing an ensemble of universes one can explain the design of the universe as a selection effect.[3] Because the ensemble contains so many universes, some are bound to be interesting and include rational agents like ourselves. That we exist and live in an apparently designed universe may surprise us, but is fully to be expected simply by chance given a large enough ensemble of universes.

Deciding whether the universe as a whole is designed always requires additional metaphysical assumptions. Thus, to reject design, Hawking needs to assume an underlying necessity to the laws of nature, Guth needs to assume "eternal inflation" (making inflation a metaphysical principle for universe construction), and Barrow and Tipler need to assume an ensemble of universes that are causally isolated from our own.[4] On the other hand, to accept design, theists like Hugh Ross begin with a prior belief in God and then formulate arguments to show how the fine-tuning of the universe actually does support design.[5] Now I don't mean to suggest there is no principled way to distinguish among metaphysical assumptions that alternately support or undercut the design of the universe. William Lane Craig and Richard Swinburne, for instance, have argued convincingly for the strength of theism as a metaphysical position in relation to cosmology.[6]

My interest in this paper, however, is different. Certainly I want to answer the question, If naturalism is false, how could we know it? But in answering this question I will consider not whether the universe as a whole is designed, but rather whether we are able to detect design within an already given universe. The universe provides a well-defined causal backdrop (many physicists think of it as a *field* characterized by *field equations*). Although one can ask whether that causal backdrop is itself designed, one can as well ask whether events and objects occurring within that backdrop exhibit features that convincingly implicate design. At issue here are two types of design: (1) the design of the universe as a whole and (2) instances of design within the universe. An analogy illustrates the difference. Consider an oil painting. An oil painting is typically painted on a canvas. One can therefore ask whether the canvas is designed. Alternatively, one can ask whether some configuration of paint on the canvas is designed. The design of the canvas corresponds to the design of the universe as a whole. The design of some configuration of paint corresponds to an instance of design within the universe.

Though not perfect, this analogy is useful. The universe is a canvas on which is depicted natural history. One can ask whether that canvas itself is designed. To decide this question, however, presupposes a serious investment in metaphysics. On the other hand, one can ask whether features of natural history depicted on that canvas are designed. In biology, for instance, one can ask whether Michael Behe's irreducibly

complex biochemical machines are designed.[7] Although metaphysical assumptions are not irrelevant here, they play a much weaker role than they do in cosmology. This is only to be expected. The origin and early evolution of the universe remain highly speculative. On the other hand, the causal backdrop for life is well-defined. The nuts and bolts of life occur at the level of biochemistry. Biochemistry provides the causal backdrop against which design in biology must be decided. That backdrop is well-defined and subject to far less speculation than the origin and early evolution of the universe.

The aim of this chapter, then, is to show how detecting design within the universe, and especially within biology, challenges naturalism. To do this, I shall review my own work on detecting design. I shall show that detecting design within the universe follows a well-defined methodology. Moreover, when applied to the irreducibly complex biochemical systems of Michael Behe, this methodology convincingly implicates design. I conclude that design within the universe renders naturalism untenable.

2. Admitting design into science

When the physics of Galileo and Newton displaced the physics of Aristotle, deterministic natural laws became the preferred mode of scientific explanation. When the physics of Bohr and Heisenberg in turn displaced the physics of Galileo and Newton, the preferred mode of scientific explanation came to include both deterministic natural laws and chance processes. Chance and necessity, to use a phrase made famous by Jacques Monod, henceforth set the boundary for scientific explanation.

That boundary needs now to be redrawn again. Chance and necessity have proven too thin an explanatory soup on which to nourish a robust science. Without invoking the rightly discarded teleologies, entelechies, and vitalisms of the past, one can still see that chance and necessity together do not span the full range of scientific explanation. A third mode of explanation is required, namely, design. Chance, necessity and design – these three modes of explanation are needed to properly span the full range of scientific explanation.

Many scientists tacitly admit this point. For instance, Richard Dawkins begins his book *The Blind Watchmaker* by stating, "Biology is the study of complicated things that give the appearance of having been designed for a purpose."[8] Statements like this echo throughout the biological literature. In *What Mad Pursuit*, Francis Crick writes, "Biologists must constantly keep in mind that what they see was not designed, but rather evolved."[9] Granted, the biological community thinks it has accounted for the apparent design in nature apart from any actual design (typically through the Darwinian mechanism of mutation and selection). The point to appreciate, however,

is that in accounting for the apparent design in nature, biologists regard themselves as having made a successful *scientific* argument against actual design.

Scientific refutation is always a double-edged sword. Claims that are refuted scientifically may be wrong, but they are not necessarily wrong. Alternatively, for a claim to be scientifically falsifiable, it must have the possibility of being true. To see this, consider what would happen if microscopic examination revealed that every cell was inscribed with the phrase "Made by Yahweh." Of course cells do not have "Made by Yahweh" inscribed on them, but that is not the point. The point is that we would not know this unless we actually looked at cells under the microscope.

Design always remains a live option in biology. A priori prohibitions against design are easily countered, especially in an age of diversity and multiculturalism where it is all too easy to ask, *Who sets the rules for science anyway?* Nonetheless, once we admit that design cannot be excluded from science on first principles, a weightier question remains: Why should we want to admit design into science? To answer this question, let us turn it around and ask instead: Why shouldn't we want to admit design into science? What's wrong with explaining something as designed by an intelligent agent? Certainly there are many everyday occurrences which we explain by appealing to design. Moreover, in our workaday lives it is absolutely crucial to distinguish accident from design. We demand answers to such questions as: Did she fall or was she pushed? Did someone die accidentally or commit suicide? Was this song conceived independently or was it plagiarized? Did someone just get lucky on the stock market or was there insider trading?

Not only do we demand answers to such questions, but entire industries are devoted to drawing the distinction between accident and design. Here we can include forensic science, intellectual property law, insurance claims investigation, cryptography, and random number generation – to name but a few. Science itself needs to draw this distinction to keep itself honest. In a recent issue of *Science*, a Medline websearch uncovered a "paper published in *Zentralblatt für Gynäkologie* in 1991 [containing] text that is almost identical to text from a paper published in 1979 in *the Journal of Maxillofacial Surgery*."[10] Plagiarism and data falsification are far more common in science than we would like to admit. What keeps these abuses in check is our ability to detect them.

If design is so readily detectable outside science, and if its detectability is one of the key factors keeping scientists honest, why should design be barred from the actual content of science? With reference to biology, why should we have to constantly remind ourselves that biology studies things that only appear to be designed, but that in fact are not designed? Isn't it at least conceivable that there could be good positive reasons for thinking biological systems are in fact designed?

The biological community's response to these questions has been to resist design resolutely. The worry is that for natural objects (unlike human artifacts), the distinction between design and non-design cannot be reliably drawn. Consider, for

instance, the following remark by Darwin in the concluding chapter of his *Origin of Species*:

> Several eminent naturalists have of late published their belief that a multitude of reputed species in each genus are not real species; but that other species are real, that is, have been independently created. . . . Nevertheless they do not pretend that they can define, or even conjecture, which are the created forms of life, and which are those produced by secondary laws. They admit variation as a *vera causa* in one case, they arbitrarily reject it in another, without assigning any distinction in the two cases."[11]

It is this worry of falsely attributing something to design (here identified with creation) only to have it overturned later that has prevented design from entering science proper.

This worry, though perhaps justified in the past, is no longer tenable. There does in fact exist a rigorous criterion for distinguishing intelligently caused objects from unintelligently caused ones. Many special sciences already use this criterion, though in a pre-theoretic form (e.g. forensic science, artificial intelligence, cryptography, archeology, and the Search for Extra-Terrestrial Intelligence). The great breakthrough of the intelligent design movement has been to isolate and make precise this criterion. Michael Behe's criterion of irreducible complexity for establishing the design of biochemical systems is a special case of this general criterion for detecting design.[12]

Before examining this criterion, I want to offer a brief clarification about the word "design." I use "design" in three distinct senses. First, I use it to denote the scientific theory that distinguishes intelligent agency from natural causes, a theory that increasingly is being referred to as "design theory" or "intelligent design theory" (IDT).[13] Second, I use "design" to denote what it is about intelligently produced objects that enables us to tell that they are intelligently produced and not simply the result of natural causes. When intelligent agents act, they leave behind a characteristic trademark or signature. The scholastics used to refer to the "vestiges of creation." The Latin *vestigium* means footprint. It was thought that God, though not physically present, left his footprints throughout creation. Hugh Ross has similarly referred to the "fingerprint of God." It is "design" in this sense – as a trademark, signature, vestige, or fingerprint – that this criterion for discriminating intelligently from unintelligently caused objects is meant to identify.[14] Lastly, I use "design" to denote intelligent agency itself. Thus, to say that something is designed is to say that an intelligent agent caused it. But note, to say that an intelligent agent caused something is not to prescribe how an intelligent agent caused it. In particular, design does not presuppose miracles.[15]

3. The complexity-specification criterion

What does the criterion for distinguishing intelligently from unintelligently caused objects look like? Although a detailed explanation and justification of this criterion is fairly technical,[16] the basic idea is straightforward and easily illustrated. Consider how the radio astronomers in the movie *Contact* detected an extra-terrestrial intelligence. This movie, based on a novel by Carl Sagan, was an enjoyable piece of propaganda for the SETI research program – the Search for Extra-Terrestrial Intelligence. To make the movie interesting, the SETI researchers actually had to find an extra-terrestrial intelligence (the non-fictional SETI program has yet to be so fortunate).

How, then, did the SETI researchers in *Contact* find an extra-terrestrial intelligence? To increase their chances of finding an extra-terrestrial intelligence, SETI researchers monitor millions of radio signals from outer space. Many natural objects in space produce radio waves (e.g., pulsars). Looking for signs of design among all these naturally produced radio signals is like looking for a needle in a haystack. To sift through the haystack, SETI researchers run the signals they monitor through computers programmed with pattern-matchers. So long as a signal does not match one of the pre-set patterns, it will pass through the pattern-matching sieve (and that even if it has an intelligent source). If, on the other hand, it does match one of these patterns, then, depending on the pattern matched, the SETI researchers may have cause for celebration.

The SETI researchers in *Contact* did find a signal worthy of celebration, namely the following:

1101110111110111111101111111111101111111111111011111111111
1111111011111111111111111111101111111111111111111111111011111
1111111111111111111111101111111111111111111111111111111110
1111111111111111111111111111111111111101111111111111111111
1111111111111111111111101111111111111111111111111111111111
1111111110111111111111111111111111111111111111111111111111
0111111111111111111111111111111111111111111111111111111011
1111111111111111111111111111111111111111111111111111111111
1101111111111111111111111111111111111111111111111111111111
1111111111110111111111111111111111111111111111111111111111
1111111111111111111111111101111111111111111111111111111111
1111111111111111111111111111111111111111111111101111111111
1111111111111111111111111111111111111111111111111111111111
1111111111101111111111111111111111111111111111111111111111
1111111111111111111111111111111111111101111111111111111111
1111111111111111111111111111111111111111111111111111111111
1111111111111101111111111111111111111111111111111111111111

11111111111111111111111111111111111111111111111111101
11111111111111111111111111111111111111111111111111111
1111111111111111111111111111111111111

In this sequence of 1126 bits, 1's correspond to beats and 0's to pauses. This sequence represents the prime numbers from 2 to 101, where a given prime number is represented by the corresponding number of beats (i.e. 1's), and the individual prime numbers are separated by pauses (i.e. 0's).

The SETI researchers in *Contact* took this signal as decisive confirmation of an extra-terrestrial intelligence. What is it about this signal that decisively implicates design? Whenever we infer design, we must establish two things – *complexity* and *specification*. Complexity ensures that the object in question is not so simple that it can readily be explained by chance. Specification ensures that this object exhibits the type of pattern that is the trademark of intelligence.

To see why complexity is crucial for inferring design, consider the following sequence of bits:

110111011111

These are the first twelve bits in the previous sequence representing the prime numbers 2, 3 and 5 respectively. Now it is a sure bet that no SETI researcher, if confronted with this twelve-bit sequence, is going to contact the science editor at the *New York Times*, hold a press conference, and announce that an extra-terrestrial intelligence has been discovered. No headline is going to read, "Aliens Master the First Three Prime Numbers!"

The problem is that this sequence is much too short (i.e. has too little complexity) to establish that an extra-terrestrial intelligence with knowledge of prime numbers produced it. A randomly beating radio source might by chance just happen to output the sequence "110111011111." A sequence of 1126 bits representing the prime numbers from 2 to 101, however, is a different story. Here the sequence is sufficiently long (i.e. has enough complexity) to confirm that an extra-terrestrial intelligence could have produced it.

Even so, complexity by itself is not enough to eliminate chance and implicate design. If I flip a coin 1000 times, I will participate in a highly complex (or what amounts to the same thing, highly improbable) event. Indeed, the sequence I end up flipping will be one in a trillion trillion trillion . . ., where the ellipsis needs twenty-two more "trillions." This sequence of coin tosses will not, however, trigger a design inference. Though complex, this sequence will not exhibit a suitable pattern. Contrast this with the previous sequence representing the prime numbers from 2 to 101. Not only is this sequence complex, but it also embodies a suitable pattern. The SETI

researcher who in the movie *Contact* discovered this sequence put it this way: "This isn't noise, this has structure."

What is a *suitable* pattern for inferring design? Not just any pattern will do. Some patterns can legitimately be employed to infer design whereas others cannot. The basic intuition underlying the distinction between patterns that alternately succeed or fail to implicate design is, however, easily motivated. Consider the case of an archer. Suppose an archer stands 50 meters from a large wall with bow and arrow in hand. The wall, let us say, is sufficiently large that the archer cannot help but hit it. Now suppose each time the archer shoots an arrow at the wall, the archer paints a target around the arrow so that the arrow sits squarely in the bull's-eye. What can be concluded from this scenario? Absolutely nothing about the archer's ability as an archer. Yes, a pattern is being matched; but it is a pattern fixed only after the arrow has been shot. The pattern is thus purely *ad hoc*.

But suppose instead the archer paints a fixed target on the wall and then shoots at it. Suppose the archer shoots a hundred arrows, and each time hits a perfect bull's-eye. What can be concluded from this second scenario? Confronted with this second scenario we are obligated to infer that here is a world-class archer, one whose shots cannot legitimately be referred to luck, but rather must be referred to the archer's skill and mastery. Skill and mastery are of course instances of design.

The type of pattern where the archer fixes a target first and then shoots at it is common to statistics, where it is known as setting a *rejection region* prior to an experiment. In statistics, if the outcome of an experiment falls within a rejection region, the chance hypothesis supposedly responsible for the outcome is rejected. Now a little reflection makes clear that a pattern need not be given prior to an event to eliminate chance and implicate design. Consider the following cipher text:

nfuijolt ju jt mjlf b xfbtfm

Initially this looks like a random sequence of letters and spaces – initially you lack any pattern for rejecting chance and inferring design.

But suppose next that someone comes along and tells you to treat this sequence as a Caesar cipher, moving each letter one notch down the alphabet. Behold, the sequence now reads,

methinks it is like a weasel

Even though the pattern is now given after the fact, it still is the right sort of pattern for eliminating chance and inferring design. In contrast to statistics, which always tries to identify its patterns before an experiment is performed, cryptanalysis must discover its patterns after the fact. In both instances, however, the patterns are suitable for

inferring design.

Patterns divide into two types, those that in the presence of complexity warrant a design inference and those that despite the presence of complexity do not warrant a design inference. The first type of pattern is called a *specification*, the second a *fabrication*. Specifications are the non-*ad hoc* patterns that can legitimately be used to eliminate chance and warrant a design inference. In contrast, fabrications are the *ad hoc* patterns that cannot legitimately be used to warrant a design inference. This distinction between specifications and fabrications can be made with full statistical rigor.[17]

4. Specification

A rigorous elucidation of the complexity-specification criterion requires explicating the dual notions of specification and complexity. We treat specification first. Specification is what must be added to an otherwise highly improbable chance event to render chance an unacceptable explanation of that event. Left to themselves highly improbable events do not warrant the elimination of chance. The precise sequence of heads and tails in a long sequence of coin tosses, the precise configuration of darts gotten by throwing darts at a dart board, and the precise seating arrangement of people at a cinema, if explained by chance, are all highly improbable events. It is only when the precise sequence of heads and tails has been recorded in advance, when the precise configuration of the darts locates all the darts at the center of the target, and when the precise seating arrangement at the cinema corresponds to seats people have been assigned on their ticket stubs – that is, when what at first blush appears to be a highly improbable event is also specified – that we are warranted in eliminating chance.

What exactly then are specifications? In the first place specifications are patterns. As I stressed earlier, what is needed to eliminate chance in explaining an event is that the event under consideration conform to a pattern. Specifications, however, are not just any old patterns, for it is always too easy to commit a *post hoc* fallacy, concocting patterns after the fact so that the only reason an event conforms to a pattern is because the pattern was read off the event, thereby leaving open the possibility that the event was due to chance after all.

Patterns thus come in two varieties, *specifications* and *fabrications*. For our purposes specifications are the good patterns, the ones that legitimately warrant eliminating chance, whereas fabrications are the bad patterns, the ones that are *ad hoc* and thus give us no warrant for eliminating chance. What then marks the difference between specifications and fabrications? An example will help answer this question. Consider the following event E, an event that to all appearances was obtained by flipping a fair coin 100 times:

THTTTHHTHHTTTTTHTHTTHHHTTHTHHHTHHHTTTTTT
THTTHTTTHHTHTTTHTHTHHTTHHHHTTTHTTHHTHTHT
HHHHTTHHTHHHHTHHHHTT E

Is E the product of chance or not? A standard trick of statistics professors with an introductory statistics class is to divide the class in two, having students in one half of the class each flip a coin 100 times, writing down the sequence of heads and tails on a slip of paper, and having students in the other half each generate purely with their minds a "random looking" string of coin tosses that mimics the tossing of a coin 100 times, also writing down the sequence of heads and tails on a slip of paper. When the students then hand in their slips of paper, it is the professor's job to sort the papers into two piles, those generated by flipping a fair coin, and those concocted in the students' heads. To the amazement of the students, the statistics professor is typically able to sort the papers with 100 per cent accuracy.

There's no mystery here. The statistics professor simply looks for a repetition of six or seven heads or tails in a row to distinguish the truly random from the pseudo-random sequences (the truly random sequences being those gotten by flipping a coin and the pseudo-random sequences being those concocted in the students' heads). In a hundred coin flips one is quite likely to see six or seven such repetitions. On the other hand, people concocting pseudo-random sequences with their minds tend to alternate between heads and tails too frequently. Whereas with a truly random sequence of coin tosses there is a 50 percent chance that one toss will differ from the next, as a matter of human psychology people expect that one toss will differ from the next around 70 percent of the time.

How then will our statistics professor fare when confronted with E? Will E be attributed to chance or to the musings of someone trying to mimic chance? According to the professor's crude randomness checker, E would be assigned to the pile of sequences presumed to be truly random, for E contains a repetition of seven tails in a row. Everything that at first blush would lead us to regard E as truly random checks out. There are exactly 50 alternations between heads and tails (as opposed to the 70 that would be expected from humans trying to mimic chance). What is more, the relative frequencies of heads and tails check out: there were 49 heads and 51 tails. Thus it is not as though the coin supposedly responsible for generating E was heavily biased in favor of one side versus the other.

Suppose, however, that our statistics professor suspects she is not up against a neophyte statistics student, but instead a fellow statistician who is trying to put one over on her. To help organize her problem, study it more carefully, and enter it into a computer, she will find it convenient to let strings of 0's and 1's represent the outcomes of coin flips, with 1 corresponding to heads and 0 to tails. In this case the following pattern D will correspond to the event E:

01000110110000010100111001011101110000000100100011
01000101011001111000100110101011110011011110111100 D

Now the mere fact that the event E conforms to the pattern D is no reason to think that E did not occur by chance, since, as things stand now, the pattern D has simply been read off of the event E.

We are back to the question of how to explain an arrow stuck in a bull's eye. Whether we are impressed with an arrow sticking in a bull's eye depends on whether the bull's eye was in place before the arrow was shot, or whether it was painted around the arrow only after the arrow landed. Where the arrow landed (= event) certainly conforms to the pattern set by the bull's eye. But if the bull's eye was painted around the arrow only after the arrow landed, there is no reason to think that where the arrow landed was dictated by something other than chance. Only if the position of the target (= pattern) is in some sense independent of whatever chance process might be responsible for the arrow's flight can this pattern preclude chance in case the arrow actually happens to be sticking in the bull's eye.

In what sense then, if any, is D independent of E? More to the point, what have we failed to consider so far that would warrant our rejecting E as the product of chance? To answer this question let us rewrite D as follows:

0
1
00
01
10
11
000
001
010
011
100
101
110
111
0000
0001
0010
0011
0100
0101
0110
0111

1000
1001
1010
1011
1100
1101
1110
1111
00 D

By viewing D in this way, anyone with the least exposure to binary arithmetic immediately recognizes that D was constructed simply by writing binary numbers in ascending order, starting with the one-digit binary numbers (i.e. 0 and 1), proceeding then to the two-digit binary numbers (i.e. 00, 01, 10 and 11), and continuing on until 100 digits were recorded. It is therefore intuitively clear that D does not describe a truly random event (i.e. an event gotten by tossing a fair coin), but rather a pseudorandom event, concocted by doing a little binary arithmetic.

Although it is now intuitively clear why E cannot properly be explained by reference to chance, we need to consider more closely why this mode of explanation fails here. We started with a putative chance event E, supposedly gotten by flipping a fair coin 100 times. Since heads and tails each have probability 1/2, and since this probability gets multiplied for each flip of the coin, it follows that the probability of E is 2^{-100}, or approximately 10^{-30}. In addition, we constructed a pattern D to which E conforms. Initially D proved insufficient to eliminate chance as the explanation of E since in its construction D was simply read off of E. Rather, to eliminate chance we had also to recognize that D could have been constructed quite easily by performing some simple arithmetic operations with binary numbers. Thus to eliminate chance we needed to employ some additional *side information*, which in this case consisted of our knowledge of binary arithmetic.

Two aspects of this side information are crucial to the elimination of chance in this example. The first is that the side information provides no information about the event E in question. Probabilistically this means that the probability of E does not change once the side information becomes available to us (a property that probabilists refer to as "conditional independence"). This is certainly the case here since our knowledge of binary arithmetic does not affect the probabilities we assign to coin tosses. The second crucial aspect of the side information is that it enables us to construct the pattern D to which E conforms. This is also the case here, where our knowledge of binary arithmetic enables us to arrange binary numbers in ascending order, and thereby construct the pattern D. The side information that enables us to eliminate chance must enable us to construct the pattern to which an event conforms, yet without recourse to the actual event. This is the crucial insight. Because the side

information provides no information about the event E (i.e. is conditionally independent of E), any pattern constructed from this side information is obtained without recourse to the event E. In this way any pattern like D that is constructed from the side information avoids the charge of being *ad hoc*.

We are now in a position to define what it means for a pattern D to be a specification of an event E, and thus for D to be capable of eliminating chance in explaining E: given a pattern D to which the event E conforms, D is a *specification* of E provided side information *K* has been identified satisfying the following two conditions

CINDE	Under the assumption that E occurred by chance, *K* provides no information about E (i.e. *K* is conditionally independent of E).
TRACT	The task of constructing D from *K* is doable (or as complexity theorists would say, "tractable").

CINDE and TRACT are respectively the *conditional independence* and *tractability* conditions for side information *K*. In case both the conditions CINDE and TRACT obtain, we say that the pattern D *specifies* the event E, and that the pattern D is a *specification*. These conditions characterize the type of side information needed to eliminate chance. In the previous example, *K* consisted of our knowledge of binary arithmetic. Our knowledge of binary arithmetic clearly does not affect the probabilities we assign to coin tosses (hence CINDE was satisfied), but does allow us to construct the pattern D representing an ascending sequence of binary numbers (hence TRACT was also satisfied). Thus it became obvious that the event E could not reasonably be attributed to chance.

As they stand, CINDE and TRACT require considerable technical elaboration. Unfortunately, I can offer only the briefest indication of what is at stake with these conditions here, and must refer the reader elsewhere for a thorough technical treatment.[18] CINDE is a probabilistic condition, stating what probabilistic properties *K* must possess if *K* is to be suitable for eliminating chance in explaining E. TRACT is a complexity-theoretic condition, stating what complexity-theoretic properties *K* must possess if *K* is to be suitable for eliminating chance in explaining E. Probabilistically, *K* must leave whatever probabilities we assign to the event E untouched. Complexity-theoretically, *K* must render the task of constructing the pattern D tractable or doable. The tasks we find doable depend on the resources we are given. Thus given current technology we find sending a person to the moon doable, but sending a person to the nearest galaxy non-doable. The doable tasks are by definition the tractable ones, the non-doable tasks the intractable ones. Complexity theory, as a generalization of computational complexity theory, ranks tasks according to difficulty, and then determines which tasks are sufficiently easy to be tractable.[19] Thus, according to TRACT the task to constructing D from *K* has to be sufficiently easy to be tractable.

5. Complexity as probability

The complexity part of the complexity-specification criterion is not the same complexity that appears in the tractability condition used to define specification. The complexity of the complexity-specification criterion is essentially a probability.[20] To see the connection between complexity and probability, consider a combination lock. The more possible combinations of the lock, the more complex the mechanism and correspondingly the more improbable that the mechanism can be opened by chance. Complexity and probability therefore vary inversely: the greater the complexity, the smaller the probability. Thus, to determine whether the complexity-specification criterion is satisfied, it is necessary to determine not only whether an object or event is specified, but also whether it has sufficiently small probability (i.e. sufficiently high complexity).

What makes a probability a small probability? To answer this question we shall need the concept of a *probabilistic resource*.[21] A probability is never a small probability in isolation, but only in relation to a set of probabilistic resources that describe the number of relevant ways an event might occur. There are two types of probabilistic resources, *replicational* and *specificational*. To understand the first of these, imagine that a massive revision of the criminal justice system has just taken place. Henceforth, a convicted criminal is sentenced to serve time in prison until he flips k heads in a row, where k is selected according to the severity of the offense (we assume that all coin flips are fair and duly recorded – no cheating is possible).

Thus for a 10-year prison sentence, if we assume the prisoner can flip a coin once every five seconds (this seems reasonable), eight hours a day, six days a week, and given that the average streak of heads is 2 ($=\Sigma_{1\leq i\leq\infty} i2^{-i}$), the prisoner will on average attempt to get a string of k heads once every 10 seconds, or 6 attempts a minute, or 360 attempts an hour, or 2880 attempts in an eight-hour work day, or 901440 attempts a year (assuming a six-day work week), or approximately 9 million attempts in 10 years. 9 million is approximately 2^{23}. Thus if we required a prisoner to flip 23 heads in a row before being released, we would on average expect to see him out within approximately 10 years. Of course specific instances will vary – some prisoners being released after only a short stay, others never recording the elusive 23 heads!

Now consider a prisoner's reaction when after approximately 10 years he finally flips 23 heads in a row. Is he shocked? Does he think a miracle has occurred? Certainly not. His reaction is more likely to be, *It's about time!* Given the number of opportunities he had for observing 23 heads in a row, he has about an even chance of getting out of prison within 10 years. There is thus nothing improbable about his getting out of prison within this span of time. It is improbable that on any given occasion when he flips his coin he will flip 23 heads in a row. But when all these occasions are considered jointly, it becomes quite probable that he will be out of prison within 10 years.

Given the event E of flipping 23 heads in a row and given n opportunities for E to occur, the question is not whether E taken by itself is improbable, but whether E remains improbable once all the opportunities for E to occur have been factored in. If we let E(n) denote the event that E occurs in at least one of these n trials, the question then is whether E(n) is improbable. Thus in the case of our coin-flipping prisoner, even though the probability of E is 2^{-23}, and therefore quite small, it is the probability of E(9,000,000)), which is approximately 1/2, and which clearly is not small, that gives the prisoner hope of actually getting out of prison within his lifetime. If the prisoner's life expectancy is better than 10 years, he stands a reasonably good chance of getting out of prison.

The following two quantities therefore offset each other: the probabilistic resources n available to the prisoner, on the one hand, and the number of heads k that he must flip in a row to secure his release, on the other. If the prisoner's probabilistic resources (=n) are big enough, it does not matter how many heads in a row (=k) he must obtain to get out of prison – his probabilistic resources will then be adequate to get him out of prison. If, on the other hand, the number of heads (=k) that must be flipped in a row is exorbitant, the prisoner's probabilistic resources (=n) will stand little chance of getting him out of prison.

Consider, for instance, a prisoner condemned to flip 100 heads in a row (=k). The probability of getting 100 heads in a row on a given trial is so small that the prisoner has no practical hope of getting out of prison, even if his life expectancy and coin tossing ability were dramatically increased. If he could, for instance, make 10 billion attempts each year to obtain 100 heads in a row (this is coin-flipping at a rate of over 500 coin flips per second, 24 hours a day, seven days a week, for a full year), then he stands only an even chance of getting out of prison in 10^{20} years. His probabilistic resources are so inadequate for obtaining the desired 100 heads that it is pointless for him to entertain hopes of freedom.

Let us now turn to the other type of probabilistic resource. In the preceding example probabilistic resources consisted of the number of opportunities for a certain event to occur. We will call this type of probabilistic resource a *replicational resource* (i.e., the number of trials or replications for a given event to occur). Replicational resources are not the only type of probabilistic resource, however. Probabilistic resources can also assume another form in which the question is not how many opportunities there are for a given event to *occur*, but rather, how many opportunities there are to *specify* an as-yet undetermined event. We wi'll call this other type of probabilistic resource a *specificational resource*. Because lotteries provide the perfect illustration of specificational resources, we consider next a lottery example.

In the interest of eliminating the federal deficit, the federal government decides to hold a national lottery in which the grand prize is the nation's gold reserves at Fort Knox. The federal government has so constructed the lottery that the probability of

any one ticket winning is 1 in 2^{100}, or approximately 1 in 10^{30}. Specifically, the lottery has been constructed so that to buy a ticket, the lottery player pays a fixed price, in this case $10, and then records a 0–1 string of length 100 – whichever string she chooses so long as it does not match a string that has already been chosen. She is permitted to purchase as many tickets as she wishes, subject only to her financial resources and the time it takes to record the 0–1 strings of length 100. The lottery is to be drawn at a special meeting of the United States Senate: in alphabetical order each senator is to flip a single coin once and record the resulting coin toss.

Suppose now that the fateful day has arrived. A trillion tickets have been sold at ten dollars a piece. To prevent cheating, Congress has enlisted the National Academy of Science. In accord with the NAS's recommendation, each ticket holder's name is duly entered onto a secure database, together with the tickets purchased and the ticket numbers (i.e. the bit strings relevant to deciding the winner). All this information is now in place. After much fanfare the senators start flipping their coins, beginning with Senator Amy Aaron and ending with Senator Zig Zygmund. As soon as Senator Zygmund has announced his toss, the database is consulted to determine whether the lottery has a winner. Lo and behold, the lottery does indeed have a winner, whose name is . . . .

From a probabilist's perspective there is one overriding implausibility to this example. The implausibility rests not with the federal government sponsoring a lottery to eliminate the national debt, nor with the choice of prize, nor with the way the lottery is decided at a special meeting of the Senate, nor even with the fantastically poor odds of anyone winning the lottery. The implausibility rests with the lottery having a winner. By all means the federal government should institute such a lottery if it seemed likely to redress the national debt, for it is obvious that if the lottery is run fairly, there will be no winner – the odds are simply too much against it. Suppose, for instance, that a trillion tickets are sold at 10 dollars apiece (this would cover the deficit as it stands at the time of this writing). An elementary calculation shows that the probability that one of these tickets (qua specifications) will match the winning string of 0's and 1's drawn by the senate cannot exceed 1 in 10^{18}. Even if we increase the number of lottery tickets sold by several orders of magnitude, there still will not be sufficiently many tickets sold for the lottery to stand a reasonable chance of having a winner.

Sometimes it is necessary to consider both types of probabilistic resources in tandem, those depending on the number of opportunities for an event to occur (i.e. replicational) as well as those depending on the number of opportunities to specify a given event (i.e. specificational). Suppose, for instance, that in the preceding lottery the Senate will hold up to a thousand drawings to determine a winner. Thus instead of having Senators Aaron through Zygmund flip their pennies in succession just once, we have them repeat this process up to a thousand times, stopping short of the thousand repetitions in case there happens to be a winner. If we now assume as before

that a trillion tickets have been sold, then for this probabilistic set-up the probabilistic resources include both a trillion specifications and a thousand possible replications. An elementary calculation now shows that the probability of this modified lottery having a winner is no greater than 1 in 10^{15}. It therefore remains highly unlikely that this modified lottery, despite the increase in probabilistic resources, will have a winner.

Probabilistic resources comprise the relevant ways an event can occur (replicational resources) and be specified (specificational resources). The important question therefore is not what is the probability of the event in question but rather what does its probability become after all the relevant probabilistic resources have been factored in? Probabilities can never be considered in isolation, but must always be referred to a relevant class of probabilistic resources. A seemingly improbable event can become quite probable when placed against a sufficiently large set of probabilistic resources. The anthropic principles that look to multiple universes bank on precisely this point: while the emergence of intelligent life in our universe is vastly improbable when considered strictly with respect to our universe, when we factor in all the possible universes that might have given rise to us, the emergence of intelligent life is rendered a virtual certainty. Whereas the status of possible universes other than our own is a matter of controversy, their role as probabilistic resources is evident.[22]

Let us now return to our original question: Given a probability p, what makes p a small probability? The primary intuition underlying small probabilities is that an event has small probability only if its probability remains small after all the probabilistic resources relevant to its occurrence have been factored in. Thus a prisoner sentenced to flip 23 heads in a row to be released from prison does not regard this event as having small probability because this event is reasonably likely given 10 years of coin flipping. A prisoner sentenced to flip 100 heads in a row to get out of prison, however, is right to regard this event as having small probability since it remains improbable even after all conceivable probabilistic resources have been factored in.

Once one has fixed all the probabilistic resources relevant to the event under consideration, how does one assess whether the probability of this event is indeed small? There seems to be an infinite regress here, for given what has been said, the probability of an event is small just in case its probability remains small after the relevant probabilistic resources have been factored in. Recall the prisoner example. A prisoner sentenced to flip k heads in a row regards this event as having small probability just in case after factoring in all his probabilistic resources for attaining this event, the recomputed probability that factors in all his probabilistic resources remains small. Determining what it is for a probability to be small is thus transformed into the problem of determining what it is for a recomputed probability to be small. As it turns out, however, this is progress, for it is much easier to say what it is for a recomputed probability to be small (i.e. a probability that has been recomputed on the basis of relevant probabilistic resources), than what it is for an arbitrary probability

to be small (i.e. a probability that is given without reference to relevant probabilistic resources).

Indeed, things could not be simpler: a probability recomputed on the basis of probabilistic resources is small just in case it is less than 1/2. The number 1/2 marks an important probabilistic boundary. Events of probability strictly less than 1/2 are less likely than not, whereas those of probability strictly greater than 1/2 are more likely than not. For an arbitrary event E of probability p, whether p is greater or less than 1/2 does not determine whether p has small probability since the probabilistic resources relevant to E's occurrence have yet to be factored in. But once those probabilistic resources are factored in, whether the probability recomputed on the basis of those probabilistic resources is greater or less than 1/2 becomes crucially important.

A prisoner sentenced to flip k heads in a row in order to be released from prison has nothing to fear from any inherent improbability of this sequence, but everything to fear from not having enough coin flips (i.e. probabilistic resources) to render this sequence more likely than not. If all the coin flips potentially available to the prisoner cannot render this sequence more likely than not, the prisoner's probability of getting out of prison will remain less than an even chance, and the prisoner should expect to remain in prison. An event whose probability remains less than 1/2, even after all the relevant probabilistic resources have been factored in, is an event which all the relevant probabilistic resources cannot conspire to render more likely than not.

To recap, a probability is never small in isolation, but only in relation to a relevant set of probabilistic resources. Once the relevant set of probabilistic resources has been prescribed, a probability is considered small just in case the probability recomputed on the basis of those resources is strictly less than 1/2. Thus for the complexity-specification criterion to be satisfied, the complexity of the event or object in question must correspond to a probability which, when recomputed against the relevant probabilistic resources, is less than 1/2.

6. Why the criterion works

Why does the complexity-specification criterion reliably detect design? To see why this criterion is exactly the right instrument for detecting design, we need to understand what it is about intelligent agents that makes them detectable in the first place. The principal characteristic of intelligent agency is choice. Whenever an intelligent agent acts, it chooses from a range of competing possibilities.

This is true not just of humans, but of animals as well as of extra-terrestrial intelligences. A rat navigating a maze must choose whether to go right or left at various points in the maze. When SETI researchers attempt to discover intelligence in the extra-terrestrial radio transmissions they are monitoring, they assume an extra-

terrestrial intelligence could have chosen any number of possible radio transmissions, and then attempt to match the transmissions they observe with certain patterns as opposed to others. Whenever a human being utters meaningful speech, a choice is made from a range of possible sound-combinations that might have been uttered. Intelligent agency always entails discrimination, choosing certain things, ruling out others.

Given this characterization of intelligent agency, the crucial question is how to recognize it. Intelligent agents act by making a choice. How then do we recognize that an intelligent agent has made a choice? A bottle of ink spills accidentally onto a sheet of paper; someone takes a fountain pen and writes a message on a sheet of paper. In both instances ink is applied to paper. In both instances one among an almost infinite set of possibilities is realized. In both instances a contingency is actualized and others are ruled out. Yet in one instance we ascribe agency, in the other chance.

What is the relevant difference? Not only do we need to observe that a contingency was actualized, but we ourselves need also to be able to specify that contingency. The contingency must conform to an independently given pattern, and we must be able independently to formulate that pattern. A random ink blot is unspecifiable; a message written with ink on paper is specifiable. Ludwig Wittgenstein in *Culture and Value* made the same point: "We tend to take the speech of a Chinese for inarticulate gurgling. Someone who understands Chinese will recognize *language* in what he hears. Similarly I often cannot discern the *humanity* in man."[23]

In hearing a Chinese utterance, someone who understands Chinese not only recognizes that one from a range of all possible utterances was actualized, but is also able to specify the utterance as coherent Chinese speech. Contrast this with someone who does not understand Chinese. In hearing a Chinese utterance, someone who does not understand Chinese also recognizes that one from a range of possible utterances was actualized, but this time, because lacking the ability to understand Chinese, is unable to specify the utterance as coherent speech.

To someone who does not understand Chinese, the utterance will appear gibberish. Gibberish – the utterance of nonsense syllables uninterpretable within any natural language – always actualizes one utterance from the range of possible utterances. Nevertheless, gibberish, by corresponding to nothing we can understand in any language, also cannot be specified. As a result, gibberish is never taken for intelligent communication, but always for what Wittgenstein calls "inarticulate gurgling."

This actualizing of one among several competing possibilities, ruling out the rest, and specifying the one that was actualized encapsulates how we recognize intelligent agency, or equivalently, how we detect design. Experimental psychologists who study animal learning and behavior have known this all along. To learn a task an animal must acquire the ability to actualize behaviors suitable for the task as well as the ability to rule out behaviors unsuitable for the task. Moreover, for a psychologist to recognize

that an animal has learned a task, it is necessary not only to observe the animal making the appropriate discrimination, but also to specify this discrimination.

Thus to recognize whether a rat has successfully learned how to traverse a maze, a psychologist must first specify which sequence of right and left turns conducts the rat out of the maze. No doubt, a rat randomly wandering a maze also discriminates a sequence of right and left turns. But by randomly wandering the maze, the rat gives no indication that it can discriminate the appropriate sequence of right and left turns for exiting the maze. Consequently, the psychologist studying the rat will have no reason to think the rat has learned how to traverse the maze.

Only if the rat executes the sequence of right and left turns specified by the psychologist will the psychologist recognize that the rat has learned how to traverse the maze. Now it is precisely the learned behaviors we regard as intelligent in animals. Hence it is no surprise that the same scheme for recognizing animal learning recurs for recognizing intelligent agency generally, to wit: actualizing one among several competing possibilities, ruling out the others, and specifying the one actualized.

Note that complexity is implicit here as well. To see this, consider again a rat traversing a maze, but now take a very simple maze in which two right turns conduct the rat out of the maze. How will a psychologist studying the rat determine whether it has learned to exit the maze? Just putting the rat in the maze will not be enough. Because the maze is so simple, the rat could by chance just happen to take two right turns, and thereby exit the maze. The psychologist will therefore be uncertain whether the rat actually learned to exit this maze, or whether the rat just got lucky.

But contrast this now with a complicated maze in which a rat must take just the right sequence of left and right turns to exit the maze. Suppose the rat must take one hundred appropriate right and left turns, and that any mistake will prevent the rat from exiting the maze. A psychologist who sees the rat take no erroneous turns and in short order exit the maze will be convinced that the rat has indeed learned how to exit the maze, and that this was not dumb luck.

This general scheme for recognizing intelligent agency is but a thinly disguised form of the complexity-specification criterion. In general, to recognize intelligent agency we must observe an actualization of one among several competing possibilities, note which possibilities were ruled out, and then be able to specify the possibility that was actualized. What is more, the competing possibilities that were ruled out must be live possibilities, and sufficiently numerous so that specifying the possibility that was actualized cannot be attributed to chance. In terms of complexity, this is just another way of saying that the range of possibilities is complex. In terms of probability, this is just another way of saying that the possibility that was actualized has small probability.

All the elements in this general scheme for recognizing intelligent agency (i.e. actualizing, ruling out, and specifying) find their counterpart in the complexity-specification criterion. It follows that this criterion formalizes what we have been

doing right along when we recognize intelligent agency. The complexity-specification criterion pinpoints what we need to be looking for when we detect design.

As a postscript it is worth pondering the etymology of the word "intelligent." The word "intelligent" derives from two Latin words, the preposition *inter*, meaning between, and the verb *lego*, meaning to choose or select. Thus according to its etymology, intelligence consists in *choosing between*. It follows that the etymology of the word "intelligent" parallels the formal analysis of intelligent agency inherent in the complexity-specification criterion.

7. Irreducible complexity

Perhaps the most compelling evidence for design in the universe comes from biology, and specifically from biochemistry. In a recent issue of *Cell*, Bruce Alberts, president of the National Academy of Sciences, remarked, "The entire cell can be viewed as a factory that contains an elaborate network of interlocking assembly lines, each of which is composed of large protein machines. . . . Why do we call the large protein assemblies that underlie cell function *machines*? Precisely because, like the machines invented by humans to deal efficiently with the macroscopic world, these protein assemblies contain highly coordinated moving parts."[24]

Nonetheless, Alberts sides with the majority of biologists in regarding the cell's marvelous complexity as only apparently designed. The Lehigh University biochemist Michael Behe disagrees. In *Darwin's Black Box* Behe presents a powerful argument for actual design in the cell. Central to his argument is his notion of *irreducible complexity*. A system is irreducibly complex if it consists of several interrelated parts so that removing even one part completely destroys the system's function. As an example of irreducible complexity Behe offers the mousetrap. A mousetrap consists of a platform, a hammer, a spring, a catch, and a holding bar. Remove any one of these five components, and it is impossible to construct a functional mousetrap.[25]

Irreducible complexity needs to be contrasted with *cumulative complexity*. A system is cumulatively complex if the components of the system can be arranged sequentially so that the successive removal of components never leads to the complete loss of function. An example of a cumulatively complex system is a city. It is possible successively to remove people and services from a city until one is down to a tiny village – all without losing the sense of community, which in this case constitutes function. From this characterization of cumulative complexity, it is clear that the Darwinian mechanism of selection and mutation can readily account for cumulative complexity. Indeed, the gradual accrual of complexity via selection mirrors the retention of function as components are successively removed from a cumulatively complex system.

But what about irreducible complexity? Can the Darwinian mechanism account

for irreducible complexity? Certainly, if selection acts with reference to a goal, it can produce irreducible complexity. Take Behe's mousetrap. Given the goal of constructing a mousetrap, one can specify a goal-directed selection process that in turn selects a platform, a hammer, a spring, a catch, and a holding bar, and at the end puts all these components together to form a functional mousetrap. Given a prespecified goal, selection has no difficulty producing irreducibly complex systems.

But the selection operating in biology is Darwinian natural selection. And this form of selection operates without goals, has neither plan nor purpose, and is wholly undirected. The great appeal of Darwin's selection mechanism was, after all, that it would eliminate teleology from biology. Yet by making selection an undirected process, Darwin drastically abridged the type of complexity biological systems could manifest. Henceforth biological systems could manifest only cumulative complexity, not irreducible complexity.

Why is this? As Behe explains in *Darwin's Black Box*:

> An irreducibly complex system cannot be produced . . . by slight, successive modifications of a precursor system, because any precursor to an irreducibly complex system that is missing a part is by definition nonfunctional. . . . Since natural selection can only choose systems that are already working, then if a biological system cannot be produced gradually it would have to arise as an integrated unit, in one fell swoop, for natural selection to have anything to act on.[26]

For an irreducibly complex system, function is attained only when all components of the system are in place simultaneously. It follows that natural selection, if it is going to produce an irreducibly complex system, has to produce it all at once or not at all. This would not be a problem if the systems in question were simple. But they're not. The irreducibly complex biochemical systems Behe considers are protein machines consisting of numerous distinct proteins, each indispensable for function, and together beyond what natural selection can muster in a single generation.

One such irreducibly complex biochemical system that Behe considers is the bacterial flagellum. The flagellum is a whip-like rotary motor that enables a bacterium to navigate through its environment. The flagellum includes an acid powered rotary engine, a stator, O-rings, bushings, and a drive shaft. The intricate machinery of this molecular motor requires approximately fifty proteins. Yet the absence of any one of these proteins results in the complete loss of motor function.[27]

The irreducible complexity of such biochemical systems counts powerfully against the Darwinian mechanism, and indeed against any naturalistic evolutionary mechanism proposed to date. Moreover, because irreducible complexity occurs at the biochemical level, there is no more fundamental level of biological analysis to which

the irreducible complexity of biochemical systems can be referred, and at which a Darwinian analysis in terms of selection and mutation can still hope for success. Undergirding biochemistry is ordinary chemistry and physics, neither of which can account for biological information. Also, whether a biochemical system is irreducibly complex is a fully empirical question: individually knock out each protein constituting a biochemical system to determine whether function is lost. If so, we are dealing with an irreducibly complex system. Mutagenesis experiments of this sort are routine in biology.[28]

The connection between Behe's notion of irreducible complexity and my complexity-specification criterion is now straightforward. The irreducibly complex systems Behe considers require numerous components specifically adapted to each other and each necessary for function. On any formal complexity-theoretic analysis, they are complex in the sense required by the complexity-specification criterion. Moreover, in virtue of their function, these systems embody patterns independent of the actual living systems (cf. section 4). Hence these systems are also specified in the sense required by the complexity-specification criterion.

Biological specification always denotes function. An organism is a functional system comprising many functional subsystems. The functionality of organisms can be cashed out in any number of ways. Arno Wouters cashes it out globally in terms of the *viability* of whole organisms.[29] Michael Behe cashes it out in terms of the *minimal function* of biochemical systems.[30] Even the staunch Darwinist Richard Dawkins will admit that life is specified functionally, cashing out functionality in terms of the *reproduction* of genes. Thus in *The Blind Watchmaker* Dawkins will write, "Complicated things have some quality, specifiable in advance, that is highly unlikely to have been acquired by random chance alone. In the case of living things, the quality that is specified in advance is . . . the ability to propagate genes in reproduction."[31]

8. So what?

There exists a reliable criterion for detecting design. This criterion detects design strictly from observational features of the world. Moreover, it belongs to probability and complexity theory, not to metaphysics and theology. And although it cannot achieve logical demonstration, it does achieve statistical justification so compelling as to demand assent. This criterion is relevant to biology. When applied to the complex, information-rich structures of biology, it detects design. In particular, the complexity-specification criterion shows that Michael Behe's irreducibly complex biochemical systems are designed.

What are we to make of these developments? Many scientists remain unconvinced. So what if we have a reliable criterion for detecting design and so what if that criterion tells us that biological systems are designed? How is looking at a biological system and

inferring it is designed any better than shrugging our shoulders and saying God did it? The fear is that design cannot help but stifle scientific inquiry.

Design is not a science-stopper. Indeed, design can foster inquiry where traditional evolutionary approaches obstruct it. Consider the term "junk DNA." Implicit in this term is the view that because the genome of an organism has been cobbled together through a long, undirected evolutionary process, the genome is a patchwork of which only limited portions are essential to the organism. Thus on an evolutionary view we expect a lot of useless DNA. If, on the other hand, organisms are designed, we expect DNA, as much as possible, to exhibit function. And indeed, the most recent findings suggest that designating DNA as "junk" merely cloaks our current lack of knowledge about function. For instance, in a recent issue of the *Journal of Theoretical Biology*, John Bodnar describes how "non-coding DNA in eukaryotic genomes encodes a language which programs organismal growth and development."[32] Design encourages scientists to look for function where evolution discourages it.

Or consider vestigial organs that later are found to have a function after all. Evolutionary biology texts often cite the human coccyx as a "vestigial structure" that hearkens back to vertebrate ancestors with tails. Yet if one looks at a recent edition of *Gray's Anatomy*, one finds that the coccyx is a crucial point of contact with muscles that attach to the pelvic floor. Now anatomy is nothing else than an exercise in design, studying the large-scale design plans/blueprints for bodies. Thus here again we find design encouraging scientists to look for function where evolution discourages it. Examples where the phrase "vestigial structure" merely cloaks our current lack of knowledge about function can be multiplied. The human appendix, formerly thought to be vestigial, is now known to be a functioning component of the immune system.[33]

Admitting design into science can only enrich science. All the tried and true tools of science will remain intact. But design also adds a new tool to the scientist's explanatory tool chest, namely, the complexity-specification criterion. Moreover, design raises a whole new set of research questions. Once we know that something is designed, we will want to know:

1. How was it constructed (cf. reverse engineering)?
2. What is its function/purpose?[34]
3. How has noise/age/friction/mutation obscured the original design plan?
4. Can we reconstruct the original design plan (cf. art restoration)?
5. What are the constraints within which the object functions well and outside of which it malfunctions or breaks?[35]
6. In what sense is the design optimal?[36]

Even so, design does much more than enrich and transform science. Indeed, design has profound metaphysical and world-view implications. The discovery of design in

the world, and especially in biology, enables us to answer the question with which we started this essay: If naturalism is false, how could we know it? Naturalism, by limiting the fundamental causation of nature to chance and necessity, must reduce all appearance of design in nature to just that – mere appearance. What the complexity-specification criterion shows, however, is that design cannot be reduced to chance and necessity. Moreover, when applied to the complex, information-rich structures of biology, like Michael Behe's irreducibly complex biochemical systems, this criterion shows that design actually is present in nature.

Naturalism, therefore, fails on its own terms. The great hope of scientific naturalism was, after all, that naturalism could be redeemed in the coin of science. That hope is now effectively dashed. There is no principled way of excluding design from the causal structure of the world. Moreover, design is, via the complexity-specification criterion, fully amenable to scientific investigation. Naturalism is therefore seen to be false on strictly scientific grounds. The logic of this conclusion is straightforward: naturalism allows only certain sorts of fundamental causes (chance and necessity). Those causes are (demonstrably) incapable of generating specified complexity. But nature exhibits specified complexity, especially in biology. Therefore naturalism is false.

If naturalism is false, how can we know it? This is how.

Notes

1 Stephen Hawking, *A Brief History of Time: From the Big Bang to Black Holes* (New York: Bantam, 1988).

2 Alan Guth, *The Inflationary Universe: The Quest for a New Theory of Cosmic Origins* (Reading, Mass.: Addison-Wesley, 1997).

3 John Barrow and Frank Tipler, *The Anthropic Cosmological Principle* (Oxford: Oxford University Press, 1986).

4 See respectively Hawking, *Brief History*, 171–5; Guth, *Inflationary Universe*, 247 ("eternal inflation"); Barrow and Tipler, *Anthropic Cosmological Principle*, 22–3.

5 Hugh Ross, *The Fingerprint of God*, 2nd ed. (Orange, Calif.: Promise, 1991), 132–8.

6 See William Lane Craig, "Design and the Cosmological Argument," in *Mere Creation: Science, Faith and Intelligent Design*, ed. William A. Dembski (Downers Grove, Ill.: InterVarsity, 1998), 335–7; Richard Swinburne, *The Existence of God* (Oxford: Oxford University Press, 1979), 107–15.

7 See Michael Behe, *Darwin's Black Box* (New York: Free Press, 1996).

8 Richard Dawkins, *The Blind Watchmaker* (New York: Norton, 1986), 1.

9 Francis Crick, *What Mad Pursuit (New* York: Basic Books, 1988), 138.

10 Eliot Marshall, "Medline Searches Turn Up Cases of Suspected Plagiarism," *Science* 279 (23 Jan 1998): 473–4.

11 Charles Darwin, *On the Origin of Species*, facsimile 1st ed. (Cambridge, Mass.: Harvard University Press, 1964 [1859]), 482.

12 Cf. Behe, *Darwin's Black Box*, 39–45.

13 Cf. Dembski (ed.), *Mere Creation*.

14 The relevant theory is developed in William A. Dembski, *The Design Inference* (Cambridge: Cambridge University Press, 1998).

15 For a non-naturalistic account of intelligent agency see Edmund Runggaldier, *Was Sind Handlungen? Eine Philosophische Auseinandersetzung mit dem Naturalismus*, in Münchener Philosophische Studien, vol. 12 (Stuttgart: Verlag W Kohlhammer, 1996).

16 For a full account see my book *The Design Inference*.

17 Ibid., ch. 5.

18 Ibid.

19 For an overview of complexity theory, see Dembski, *The Design Inference*, ch. 4.

20 For probability as a special case of complexity, see Dembski, *The Design Inference*, sec. 4.6.

21 For a full account of probabilistic resources, see Dembski, *The Design Inference*, ch. 6.

22 John Leslie analyzes this line of reasoning masterfully in his book *Universes* (London: Routledge, 1989).

23 Ludwig Wittgenstein, *Culture and Value*, ed. G. H. von Wright, trans. P. Winch (Chicago: University of Chicago Press, 1980), 1e.

24 Bruce Alberts, "The Cell as a Collection of Protein Machines: Preparing the Next Generation of Molecular Biologists," *Cell* 92 (8 Feb 1998): 291.

25 Behe, *Darwin's Black Box*, 39–45.

26 Ibid., 39.

27 Ibid., 69–72.

28 See, for example, Nicholas Gaiano, Adam Amsterdam, Koichi Kawakami, Migeul Allende, Thomas Becker, and Nancy Hopkins, "Insertional Mutagenesis and Rapid Cloning of Essential Genes in Zebrafish," *Nature* 383 (1996): 829–32; Carolyn K. Suzuki, Kitaru Suda, Nan Wang, and Gottfried Schatz, "Requirement for the Yeast Gene *LON* in Intramitochondrial Proteolysis and Maintenance of Respiration," *Science* 264 (1994): 273–6; Qun-Yong Zhou, Carol J. Qualfe, and Richard D. Palmiter, "Targeted Disruption of the Tyrosine Hydroxylase Gene Reveals that Catecholamines are Required for Mouse Fetal Development," *Nature* 374 (1995): 640–3.

29 Arno Wouters, "Viability Explanation," *Biology and Philosophy* 10 (1995): 435–57.

30 Behe, *Darwin's Black Box*, 45–6.

31 Dawkins, *Blind Watchmaker*, 9.

32 John W. Bodnar, Jeffrey Killian, Michael Nagle, and Suneil Ramchandani, "Deciphering the Language of the Genome," *Journal of Theoretical Biology* 189 (1997): 183.

33 Percival Davis and Dean Kenyon, *Of Pandas and People*, 2nd ed. (Dallas, Tex.: Haughton, 1993), 128.

34 Note that we can detect design without knowing what something was designed for. There is a room at the Smithsonian filled with obviously designed objects for which no one has a clue

about their purpose. See Del Ratzsch's article "Design, Chance, and Theistic Evolution," in Dembski, ed., *Mere Creation*, 294.

35 Design implies constraints. An object that is designed functions within certain design constraints. Transgress those constraints and the object functions poorly or breaks. Moreover, we can discover those constraints empirically by seeing what does and doesn't work. This simple insight has tremendous implications not just for science but also for ethics. If humans are in fact designed, then we can expect psychosocial constraints to be hardwired into us. Transgress those constraints, and we personally as well as our society will suffer. There's plenty of empirical evidence to suggest that many of the attitudes and behaviors our society promotes undermine human flourishing. Design promises to reinvigorate that ethical stream running from Aristotle through Aquinas known as natural law. See J. Budziszewski, *Written on the Heart: The Case for Natural Law* (Downers Grove, Ill.: InterVarsity, 1997).

36 Optimality of design is a separate question from design. Something can be designed without being optimally designed. What's more, optimality of design must not be confused with perfection of design. Indeed, there is no such thing as "perfect design." As Henry Petroski aptly notes, "All design involves conflicting objectives and hence compromise, and the best designs will always be those that come up with the best compromise." See Henry Petroski, *Invention by Design: How Engineers Get from Thought to Thing* (Cambridge, Mass.: Harvard University Press, 1997), 30.

Index

Aaron, Senator Amy 268
'abstract particulars' 70
Akaike criterion 59, 60, 61
Alberts, Bruce 273
A- and B-theories of time 232–4
anti-foundationalism 42
anti-naturalism 43
antirealism 13
Aquinas, Thomas 93, 244
Arguments-on-Paper Thesis, the 42
Aristotle 3, 120, 193, 215, 231, 255
Armstrong, D.M. 69, 77, 78–9, 83, 84, 96–104, 156–7
'aspects' 70
Atkins, Peter 236
atomic theory of matter 76
attribute-agreement 70
Aufbau projects 43
Augustine, St 194
Aune, Bruce 77
Austin, J.L. 153

Barrow, John 217, 226, 228, 254
Bedau, Mark 125
Behe, Michael 255, 257, 275, 277
  Darwin's Black Box 273, 274
Bekenstein–Hawking formula 229
belief 32–3
Bergmann, Gustav 83
Bhaskar, Roy 73
bio-social optimism 199
Block, Ned 140, 141
Bodnar, John 276
Bohr, Niels 255
Bondi, Hermann 219
BonJour, Laurence 79
Borde, Arvind 224, 225, 231
Boyle, Joseph 197
Bradley, F.H. 97, 103
Braine, David 173
Brandenberger, Robert 224
Brentano, F. 40
Bunge, Mario 235

Campbell, C.A. 143
Campbell, Donald 199–200
Campbell, Keith 69, 77, 100, 102, 103
  early version nominalism 81–5
  revised version of nominalism 86–96
Cartesian dualism 5, 79
'cases' 70
causal theories of action 159–61
Chalmers, David 157, 175, 177, 178, 180
chaotic inflationary model 224–5
Chomsky, Noam 147
Chu Hsi 193
Churchland, Paul 8, 50, 133, 135–8, 141–4, 146–7
Clarke, Randolph 164, 165
Coherentism 43
complexity as probability 265–70
complexity-specification criterion 258–61
  detection of design 270–3
conceptualism 113
Council, Richard 101
consciousness 38, 148–9
conservatism, epistemological principle of 20–1
Contact (film) 258–9
contemporary materialism 156

contingent identity 6
conventionalism 113
Copernicus 51
Copleston, Frederick 215
core methodological naturalism 10
core ontological naturalism 10
Core Scientism 10–14, 16–17, 19, 22
correspondence 35–6
  naivety of 36–9
Craig, William Lane 254
Crick, Francis: *What Mad Pursuit* 255
critical thinking 192
cumulative complexity 273
curved space, theory of 51

Darwin, Charles 76, 77, 274
  Origin of Species 257
Davidson, Donald 8, 164
Davies, P.C.W. 217, 231
Dawkins, Richard: *Blind Watchmaker, The* 125, 255, 275
De Vito, Scott 60
Democritus 62
Dennett, Daniel 8, 40–1, 50, 77, 134, 147, 149
Derrida, Jacques 36, 38
Descartes, René 5, 6, 24, 25, 173, 174, 204
  Discourse on Method 43
descriptivism 27
design theory 257
Dewey, John 24, 27, 30
Dirac, Paul 51
divine attributes 150–1
Doe, Jane 138
Dostoyevsky, F.
  Brothers Karamazov 208
Dretscke 53
dualism 134, 135, 157
Ducasse, C.J. 143

Eddington, Sir Arthur 217
Einstein, Albert 44, 216, 227
  theory of gravitation 51
electron, theory of (Dirac) 51
electroweak theory 51
eliminative materialism 135–8
eliminative methodological naturalism 9
eliminative ontological naturalism 8
emergentism 180
empirical sciences 5, 8, 10, 11
Epicurus 62, 235
epistemic warrant 21, 22
epistemological naturalism 20
ethical intuitionism 27
evolution, theory of 76
evolutionary biology 76
Ewing, A.C. 143
externalism 44, 74, 75
extreme nominalism 70
extreme particularism 197–8

falsity 35
Fichte, I.H. von 37
Field, Hartry 68
Finnis, John 197
first philosophy 3–4, 69
Fodor, Jerry 9, 54
Forster, Malcolm
  on simplicity 58–61
Foster, John 141
foundationalism 43, 92
Frankfurt, Harry 159, 160
Frege, Gottlob 25, 41
Friedman, Alexander 216
Friedman–Lemaître (Big Bang model) 216–18
full-orbed naturalism 62

Galileo 255
Gamow, George 220
Gauss, J.K.F. 51
General Theory of Relativity 51, 216, 231
Gibbard, Allan 200
Gilligan, Carol 197
Ginet, Carl 158, 165, 166–7
Goetz, Stewart 143
Gold, Thomas 219
Goldman 30
Gott, J. Richard 231, 232
Grand Story 75–9
gravitation, theory of 51
Green, T.H. 24
Gribbin, John 235
Griffin, James 196
Grisez, Germain 197

Grossmann, Reinhardt 68, 70
group theory 51
Grünbaum, Adolf 238–9
Guth, Alan 254

Habermas, J. 191
Hacker, P.N.S. 152
Hansen, Norwood 42
Hare, R.M. 191, 192, 193
Hart, W.D. 143
Hartle, James 225
Hartle–Hawking models 226
Hawking, Stephen 220, 225, 227, 228, 253
Hawking–Penrose Singularity Theorems 220
Hegel, G.W.F. 36, 37, 198
Heil, John 157, 171, 175
Heisenberg, W.K. 255
heliocentric model (Copernicus) 51
Hicks, G. Dawes 90
Hook, Sidney 24, 30
Horkheimer 42
Hoyle, Sir Fred 217, 218
Hubble, Edwin 216
Hull, David 77
Hume, David 32, 37, 81, 133, 215, 230, 244
Husserl, Edmund 24, 25, 41, 43, 71, 98

identity materialism 139–47
integrative dualism 152–3
intelligent design theory (IDT) 257
intentionality 50
internal symmetry, theory of 51
internalism 44, 74–5
intuitive thinking 192
Inwagen, Peter van 116
irreducible complexity 273–5
Isham, Christopher 222, 223, 236

Kagan, Shelly, 195
  Limits of Morality, The 194
Kane, Robert 162–4
Kanitscheider, Bernulf 235, 237
Kant, I. 24, 25, 33, 36, 37
  on morality 189–90, 191–2, 193, 197, 198, 201–8
K-containing regions 116
Kenny, Anthony 152
Kicher, Patricia 30
Kieseppä, I.A. 60
Kim, Jaegwon 26, 77, 87, 134, 170–4, 177–9
kind concepts 115–16
kind-constituting matter 116
knowledge, definition 31–4
Kornblith 42
Kukla, Andre 60

Leibniz, G.W. 142, 215, 231
Lemaître, Georges 216
Lewis, David 56, 240, 241–2
Lewis, H.D. 143
libertarian freedom 161
Linde, Andrei 224, 225
linguistic rule theories 44
Li-Xin Li 231, 232
local naturalists 78
Locke, John 36, 37
logical grounding 33
logical psychologism 43
logical relations 41–2
Lotze, R.H. 25
Lucretius 62, 235
Luther, Martin 194, 202
Lycan, William 141, 144, 147
  Conscious and Experience 134

Mackie, J.L. 133, 135, 152
  Miracle of Theism, The 133
Maddy, Penelope 68–9
Malcolm, Norman 215
Marx, K. 198
material constitution, problem of 114
materialism 4–5, 6, 29
Mavrodes, George 135
Mayr, E. 76, 77
McGinn, Colin 68, 139, 140, 147, 148, 150, 151, 153, 176
mental causation, nature of 151–2
metaphilosophical naturalism 62
metaphysical cosmology 225
metaphysics and epistemology 143–5
methodological monism 28

methodological naturalism 9, 14–17
  conceptual necessity 17–18
  metaphysical necessity 18–19
Midas touch picture of consciousness 38
Mijuskovic, B. 151
Mill, John Stuart 191, 192, 194
Millikan, Ruth Garrett 53, 54, 56, 124, 125
Milton, John: *Paradise Lost* 35
mind–body identity 5–6, 134, 151, 169
mind–brain relation 176–7
mind-dependent material objects 118–19
modal antirealism 113–14
'moments' 70
monism 25
Monod, Jacques 255
Moore, G.E. 27
moral demand 190–1
moral faith 201–8
  providence 203–6
  pure rationalism 201–3
  experience of evil 206–8
moral gap 189
Murphy, A.E. 30

Nagel 140, 173
National Academy of Science 268
National Endowment for the Humanities 10
National Science Foundation 10
natural capacities 191–3
natural selection 274
naturalism
  definition 14–17, 25–7
  logical status of 14–19
  syntactic tautology 14
  traditional realist properties and 79–80
naturalist ontology 78
naturalistic monism 30
naturalized epistemology 26
Neander, Karen 124–7
Newton, I 255
  theory of gravitation 51
Nietzsche, F.W. 193, 201, 207
Noddings, Nel 197, 198
noetic unity 34, 42–4
nominalism 67, 70
non-eliminative non-reductive methodological naturalism 9
non-eliminative non-reductive ontological naturalism 8
non-eliminative reductive ontological naturalism 8
non-eliminative reductive methodological naturalism 9
nonreductive materialism 147–53
normativity 26–7
Novikov, I.D. 221

O'Connor, Timothy 165, 179–80
O'Shaughnessy, Brian 150
ontological dualism 5
ontological naturalism 4–6
  definition 50
ontological pluralism 5
oscillating models 220–2
overfitting 59

Papineau, David 8, 53, 56, 74, 78, 156–7
Parfit, Derek 215
Penrose, Roger 220, 229
'perfect particulars' 70
phenomenalism 37
philosophical monism 3
philosophical naturalism 3
philosophical Puritanism 30
physicalism 24, 29–30, 34, 68
physics-ism 29–30
Plantinga, Alvin 43, 122, 122
Plato 3, 24, 25, 67, 99
Pluralism 30
Poland, Jeffrey 78
Popper, Karl 42
positivism 37
possible holy being 193–4
Post, John 237
pragmatism 13, 61
Principle of In toto Location (PIL) 97–8
Principle of Instantiation 99
probabilistic resource 266
proper-function-phenomena 119–20, 122

properties 67–104
  attribute agreement 70
  nominalism 70
  realist view 70–2
'property-instances' 70
Ptolemaic system 51
Putnam, Hilary 24, 26, 36, 38–40
Putnam–Kripke–Marcus theory of linguistic reference to natural kinds 19

qualia 145–6
quantum gravity models 225–30
quasi-tropes 86, 90
Quine, W.V. 6–7, 8, 9, 17, 25–7, 30, 39, 42, 44, 62, 68
  epistemiological naturalism of 19–22

Ramsey, Frank 56
Rawls, John 191, 193
realism 70–1
red-shift 216
referential opacity 141–2
Reichenbach, Hans 61
Relativity, General Theory of 51, 216, 231
reliability 53–5, 61
replicational resource 266, 267
representation, theories of 36
representational naturalism 50
  centrality of reliability to 53–5
Rey, Georges 141, 147, 149
Richardson, R.C. 151
Riemann, G.F.B. 51
Robinson, Howard 67
Rorty 38, 39
Ross, Hugh 254, 257
Russell, Bertrand 215

Sagan, Carl 258
Santayan, George 24
scientific naturalism, contemporary 72–9
  naturalist epistemic attitude 73–5
  naturalist grand story 75–9
scientific realism 56–8, 74
  definition 49–50
scientism 29–30
Scotus, Duns 86, 96
Search for Extra-Terrestrial Intelligence (SETI) response 258–9, 270
Searle, John 28, 23, 73, 76, 78, 147, 169–75, 177, 179–81
Seibt, Johanna 72, 102
Sellars, Roy Wood 32
Sellars, Wilfrid 68, 69, 73, 78, 110
Silk, Joseph 221
simplicity 50–3, 58–61
Singularity Theorems 220
Slipher, Vesto 216
Smart, J.J.C. 8, 9
Smith, Quentin 218, 235, 238, 240, 241–3
Sober, Elliott on simplicity 58–61
social constructionism 43–4
Socrates 112
Sophocles 206, 207
space (space–time) trope 91
spatial extension 4
specification 261–5
specificational resource 266
Spener, Philipp Jakob 202
statistical normality (SN) accounts of proper functions 122, 123–4
Stebbing, L.S. 43
stereotopic fallacy 144
Stich, Stephen 30, 50
Stout, G.F. 90
Strawson, Galen 140, 147
strong naturalists 78
Stubenberg, Leopold 140, 147
Suarez, Francis 86, 92
supervenience naturalism 168, 170, 180
Swinburne, Richard 143, 151, 152, 234, 254

Tarski, Alfred 21
Taylor, Richard 162
teleological properties 53
theism 148–50
Thomists 197
Thomson, Judith Jarvis 196
Thucydides 206, 207
Tipler, Frank 217, 254
'tropes' 70, 81–5
truth 33, 34–6, 39–41
Tryon, Edward 222
Turney, Peter 61

'unit properties' 70
unity of apperception 44
universe, origin of, theories 215–30
  chaotic inflationary model 224–5
  oscillating models 220–2
  quantum gravity models 225–30
  standard Big Bang model (Friedman–Lemaître model) 216–18
  steady state model 218–31
  vacuum fluctuation models 222–3, 237
utilitarianism 191

vacuum fluctuation models 222–3, 237
verificationism 43
Vilenkin, Alexander 224, 225, 227, 231
Vilenkin models 226
vitalism 145

Wagner, Steven 73
Warner, Richard 73
weak naturalists 78
Weinberg, Steven 51, 52
Wesel, Elie 208
Weyl Curvature Hypothesis 229
Wheeler, John 216
Williams, Bernard 206, 207
Williams, D.C. 83, 95, 96
Wittgenstein, Ludwig 215
  Culture and Value 271
Wundt, Wilhelm 43

Yates, John 240

Zeldovich, Ya. B. 221
Zygmund, Senator Zig 268